

Weak KAM theory for discount Hamilton-Jacobi equations and its application

Hiro Yoshi Mitake ^{*} and Kohei Soga [†]

Abstract

Weak KAM theory for discount Hamilton-Jacobi equations and corresponding discount Lagrangian/Hamiltonian dynamics is developed. Then it is applied to error estimates for viscosity solutions in the vanishing discount process. The main feature is to introduce and investigate the family of α -limit points of minimizing curves, with some details in terms of minimizing measures. In error estimates, the family of α -limit points is effectively exploited with properties of the corresponding dynamical systems.

Keywords: discount Hamilton-Jacobi equation; Lagrangian dynamics; Hamiltonian dynamics; weak KAM theory; α -limit point; error estimate

AMS subject classifications: 35B40, 37J50, 49L25

1 Introduction

Weak KAM theory states the connection between viscosity solutions of the Hamilton-Jacobi equation with constant $c \in \mathbb{R}^n$ and $h(c) \in \mathbb{R}$,

$$(1.1) \quad H(x, c + v_x(x)) = h(c) \quad \text{in } \mathbb{T}^n$$

and the corresponding Hamiltonian (resp., Lagrangian) dynamics generated by H (resp., the Legendre transform L of H). The function H is assumed to be a Tonelli Hamiltonian, i.e.,

$$(H1) \quad H(x, p) : \mathbb{T}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad C^2,$$

$$(H2) \quad H_{pp} \text{ is positive definite,}$$

$$(H3) \quad \lim_{|p| \rightarrow +\infty} \frac{H(x, p)}{|p|} = +\infty \text{ uniformly.}$$

^{*}Institute of Engineering, Division of Electrical, Systems and Mathematical Engineering, Hiroshima University, 1-4-1 Kagamiyama, Higashi-Hiroshima, 739-8527, Japan (hiroyoshi-mitake@hiroshima-u.ac.jp) Partially supported by JSPS grants: KAKENHI #15K17574, #26287024, #16H03948.

[†]Department of Mathematics, Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, 223-8522, Japan (soga@math.keio.ac.jp). Partially supported by JSPS Grant-in-aid for Young Scientists (B) #15K21369.

It is well-known that for each $c \in \mathbb{R}^n$ there exists the unique constant $h(c)$ for which (1.1) admits a viscosity solution, where such a viscosity solution is not unique with respect to c even up to constants. The function $c \mapsto h(c)$ is called the *effective Hamiltonian*.

There are several techniques to construct or approximate viscosity solutions of (1.1) such as the vanishing viscosity method, a finite difference approximation and a discount approximation. In this paper, we consider a discount approximation of the form

$$(1.2) \quad \varepsilon v^\varepsilon + H(x, c + v_x^\varepsilon) = h(c) \quad \text{in } \mathbb{T}^n,$$

where $h(c)$ is same as the effective Hamiltonian in (1.1) and $\varepsilon > 0$. The problem (1.2) is uniquely solvable for each c because of the term “ $\varepsilon v^\varepsilon$ ”, which is sometimes called a *discount factor* in the theory of optimal control. Recently it is proved that there exists a viscosity solution v^* of (1.1) such that $v^\varepsilon \rightarrow v^*$ uniformly as $\varepsilon \rightarrow 0+$ based on weak KAM theory [7] and on the nonlinear adjoint method [22] ([22] covers some degenerate elliptic problems). Note that this convergence result holds even in the case where (1.1) has more than one solution beyond constant difference. Thus, this convergence itself is highly nontrivial, and is sometimes called the *selection problem* in Hamilton-Jacobi equations. The selection criterion in the vanishing discount process is given in [7], [22], but a rate of convergence is still an interesting open problem.

We will give partial results on a rate of convergence, investigating weak KAM theory to discount Hamilton-Jacobi equations and the corresponding dynamical systems. The corresponding dynamical systems are Lagrangian/Hamiltonian dynamics with a friction term (3.2)/(3.3), where they are still equivalent through the Legendre transform but there is not the Hamiltonian structure due to the friction term. We call the dynamical systems the *discount Lagrangian/Hamiltonian dynamics*. Variational characterization of v^ε by discount value functions is available as (3.1). Each minimizing curve of (3.1) yields an orbit of the discount Lagrangian/Hamiltonian dynamics. Then, discussion similar to standard weak KAM theory is available to discount Hamilton-Jacobi equations. Our investigation owes to the fact that discount Hamilton-Jacobi equations are of the class that admits the method of characteristics with deterministic dynamical systems. This is not the case with the vanishing viscosity method or a finite difference method, which makes the problem more difficult in general.

One of main features of this paper is to focus on α -limit points of minimizing curves to recover weak KAM theory, which does not seem to be done yet in the literature. Let us first recall original *Mather's minimizing problem*:

(M) *Consider*

$$(1.3) \quad \inf \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, \xi) - c \cdot \xi \, d\mu = -h(c),$$

where the infimum is taken over all probability measures on $\mathbb{T}^n \times \mathbb{R}^n$ which are invariant under the Euler-Lagrange flow of (2.2).

A minimizing measure of (M) is called a *Mather measure*, and the union of the supports of all Mather measures with each c is called the *Mather set*. There are many ways to construct a Mather measure, e.g., [21], [12], [20], [11], [22]. General properties of

the Mather set in regards to viscosity solutions are well-known. It is also known that each minimizing curve of a viscosity solution to (1.1) induces a Mather measure, whose support is contained in the family of α -limit points of the curve [12], [10]. Furthermore, Poincaré's recurrence theorem states that, for each Mather measure μ , μ -a.e. points of $\text{supp}(\mu)$ are recurrence. Hence one can expect inclusion between the Mather set and the family of α -limit points of all minimizing curves of all viscosity solutions to (1.1). However, since α -limit points of minimizing curves are obtained viscosity-solution-wise, it is not trivial to find a general property of the family independently from a choice of a viscosity solution.

Our first aim is to specify the family of all α -limit points of all minimizing curves for each c to be one of the main objects in weak KAM theory such as the Mather set, Aubry set, Mañé set, etc. We show that the family of α -limit points contains the Mather set and is contained in the Aubry set, summarizing its general properties.

Our second aim is to generalize what we will observe in the non-discount problem to the discount problem, where complementary analysis of viscosity solutions and the corresponding dynamical systems with minimizing curves is effectively done. Parallel argument with minimizing measures is also available. Indeed, we consider a minimizing problem associated with the discount problem:

(M1) $^\varepsilon$ *Let v^ε be the (unique) viscosity solution of (1.2). Consider*

$$\inf \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, \xi) - c \cdot \xi - \varepsilon v^\varepsilon(x) d\mu = -h(c),$$

where the infimum is taken over all probability measures on $\mathbb{T}^n \times \mathbb{R}^n$ which are invariant under the flow of the discount Euler-Lagrange equation (3.2).

This minimizing problem can be considered as a natural generalization of (M). The minimizing measures of (M1) $^\varepsilon$ are, by definition, invariant under the flow of the discount Lagrangian dynamics, and the union of their support is contained in the family of α -limit points of all minimizing curves of the viscosity solution to (1.2). We prove that the family of α -limit points is obtained with properties similar to those in the non-discount problem. We also compare (M1) $^\varepsilon$ with another type of a generalization of (M) to the discount problem:

(M2) $^\varepsilon$ *Let v^ε be the viscosity solution of (1.2). Consider for each $x_0 \in \mathbb{T}^n$,*

$$\inf \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, \xi) - c \cdot \xi d\mu = -h(c) + \varepsilon v^\varepsilon(x_0),$$

where the infimum is taken over all probability measures on $\mathbb{T}^n \times \mathbb{R}^n$ which satisfy

$$(1.4) \quad \int_{\mathbb{T}^n \times \mathbb{R}^n} \varphi_x(x) \cdot \xi d\mu = \varepsilon \varphi(x_0) - \varepsilon \int_{\mathbb{T}^n \times \mathbb{R}^n} \varphi(x) d\mu \quad \text{for all } \varphi \in C^1(\mathbb{T}^n, \mathbb{R}).$$

This problem is first introduced in [14], and then the selection problem in the vanishing discount process is solved with its minimizing measures [7], [22]. Minimizing measures of

$(M2)^\varepsilon$ are not invariant, and their support has no information on α -limit points in general. Regardless of such difference between $(M1)^\varepsilon$ and $(M2)^\varepsilon$, both minimizing measures tend to some of minimizing measures of (M) , i.e., Mather measures, as the discount parameter ε goes to zero.

Our third aim is to apply the analysis on the family of α -limit points of minimizing curves to an error estimate between v^ε and the unique limit v^* of the vanishing discount process obtained in [7], [22]. We take the following strategy to obtain an estimate:

- (i) First, we estimate an error on the set of the family of α -limit points.
- (ii) Second, we estimate the time for each point of \mathbb{T}^n to fall into the family of α -limit points along a minimizing curve of v^ε and v^* .

This strategy would work well, if properties of the corresponding dynamical systems are a priori known. In this paper, we show two successful examples.

The first example is a simple one-dimensional problem whose corresponding Lagrangian dynamics and Hamiltonian dynamics possesses hyperbolicity, where exponential asymptotics toward the family of α -limit points is available. Since hyperbolicity is persistent for any small (non-Hamiltonian) perturbation, the discount Lagrangian/Hamiltonian dynamics can be still studied by means of hyperbolicity. In this discussion, we specify the families of α -limit points of minimizing curves of v^ε and v^* , where we can also observe which Mather measures are obtained as the limit of minimizing measures of $(M1)^\varepsilon$ through the vanishing discount process (selection of Mather measures). It is interesting to note that not every Mather measure is available. The difference between the selection criterion of the vanishing discount process and that of the vanishing viscosity process is also visible. See Theorem 4.1 and Remark 4.2 for details.

The second example is that an exact viscosity solution admits a KAM torus. In this case, the family of α -limit points is equal to the whole set \mathbb{T}^n and each minimizing curve is ergodic on \mathbb{T}^n with a Diophantine rotation vector. We obtain an error estimate using a rate of ergodicity given by the Diophantine exponent.

Although we demonstrate error estimates only in two special cases as a first step, our results imply that the error between v^ε and v^* may depend highly on dynamics of the corresponding dynamical systems in general.

Finally we refer to recent development of analysis on the selection problem related to (1.1). A generalization of [7], [22] to second-order fully nonlinear problems is given in [17] introducing a dual method. A discrete version of [7] is shown in [8]. Some non-convex cases are studied in [16]. A partial result to the selection problem in the vanishing viscosity method is obtained in [3], [2], and that in a finite difference method is given in [24]. The lecture note [19] states lots about related topics based on the nonlinear adjoint method. Except for the case of the vanishing discount approximation in a convex setting, the selection problem on (1.1) is still rather open.

The paper is organized as follows: In Section 2, we recall some of weak KAM theory for the non-discount problem and investigate the family of α -limit points. In Section 3, we extend the argument in Section 2 to the discount problem, with details on $(M1)^\varepsilon$ and $(M2)^\varepsilon$. In Section 4, we demonstrate error estimates for v^ε and v^* in the above two cases.

Acknowledgement. The authors would like to thank Professor Albert Fathi for valuable comments on this work. The authors are grateful to Professors Diogo A. Gomes and Yifeng Yu for valuable comments for the previous version of the manuscript.

2 Weak KAM theory for non-discount problem

We overview weak KAM theory and investigate the family of α -limit points of minimizing curves, which is denoted by $\mathcal{M}_\alpha(c)$. We show that $\mathcal{M}_\alpha(c)$ is a set between the (projected) Mather set $\mathcal{M}(c)$ and the (projected) Aubry set $\mathcal{A}(c)$.

2.1 Viscosity solution and minimizing curve

In this section we recall several known facts on viscosity solutions of (1.1) (see, e.g., [6] for more details). Let $L(x, \xi)$ be the Legendre transform of $H(x, p)$, which satisfies under (H1)–(H3),

$$L(x, \xi) = \sup_{p \in \mathbb{R}^n} \{\xi \cdot p - H(x, p)\}.$$

Let $v(x)$ be a viscosity solution of (1.1). Note that v is Lipschitz continuous. Then $v(x)$ satisfies for each $x \in \mathbb{T}^n$ and $T > 0$,

$$(2.1) \quad v(x) = \inf_{\gamma \in AC, \gamma(0)=x} \left\{ \int_{-T}^0 (L(\gamma(s), \gamma'(s)) - c \cdot \gamma'(s) + h(c)) ds + v(\gamma(-T)) \right\},$$

where AC is the family of all absolutely continuous curves $[-T, 0] \rightarrow \mathbb{T}^n$. By Tonelli's theory, one can find at least one minimizing curve γ^* of the variational problem (2.1), which is a C^2 -solution of the Euler-Lagrange equation generated by $L^c := L - c \cdot \xi$,

$$(2.2) \quad \frac{d}{ds} \{L_\xi^c(x(s), x'(s))\} = L_x^c(x(s), x'(s)) \Leftrightarrow \frac{d}{ds} \{L_\xi(x(s), x'(s))\} = L_x(x(s), x'(s)).$$

Let ϕ_L^s denote the Euler-Lagrange flow of (2.2), i.e., $\phi_L^s(x(0), x'(0)) = (x(s), x'(s))$. The viscosity solution v is differentiable on the above minimizing curve γ^* satisfying for $s \in [-T, 0)$,

$$(2.3) \quad v_x(\gamma^*(s)) = L_\xi^c(\gamma^*(s), \gamma^{*'}(s)) = L_\xi(\gamma^*(s), \gamma^{*'}(s)) - c.$$

In particular, if $v_x(x)$ exists, (2.3) holds for $s = 0$ and γ^* is the unique minimizing curve for $v(x)$. The minimizing curve γ^* can be extended to $(-\infty, 0]$. By the variational property, we obtain for any $\tau > 0$ and $\tilde{\tau} \geq \tau$,

$$\begin{aligned} v(\gamma^*(-\tau)) &= \int_{-\tilde{\tau}}^{-\tau} (L(\gamma^*(s), \gamma^{*'}(s)) - c \cdot \gamma^{*'}(s) + h(c)) ds + v(\gamma^*(-\tilde{\tau})) \\ &= \int_{-\tilde{\tau}+\tau}^0 (L(\gamma^*(s-\tau), \gamma^{*'}(s-\tau)) - c \cdot \gamma^{*'}(s-\tau) + h(c)) ds + v(\gamma^*(-\tilde{\tau})), \\ v(x) &= \int_{-\tau}^0 (L(\gamma^*(s), \gamma^{*'}(s)) - c \cdot \gamma^{*'}(s) + h(c)) ds + v(\gamma^*(-\tau)), \end{aligned}$$

and hence we see that (2.3) holds for all $s < 0$.

We call an extended minimizing curve defined on $(-\infty, 0]$ a one-sided global minimizing curve or just a minimizing curve. Since $\gamma^*(s)$ solves (2.2) for $s \leq 0$, we see that $(\gamma^*(s), p^*(s))$ with $p^*(s) := L_x(\gamma^*(s), \gamma^{*'}(s))$ is a C^1 -solution of the Hamiltonian system,

$$(2.4) \quad \begin{cases} x'(s) = H_p(x(s), p(s)), \\ p'(s) = -H_x(x(s), p(s)). \end{cases}$$

Let ϕ_H^s denote the Hamiltonian flow of (2.4). Due to equivalence between (2.2) and (2.4), and (2.3), we have $p^*(s) = c + v_x(\gamma^*(s))$ for all $s < 0$, which means that for all $s < 0$,

$$(\gamma^*(s), p^*(s)) \in \text{graph}(c + v_x) := \{(x, c + v_x(x)) \mid x \in \mathbb{T}^n \text{ such that } v_x(x) \text{ exists}\}.$$

Therefore we have

Theorem 2.1. *The set $\text{graph}(c + v_x)$ is backward invariant under the Hamiltonian flow ϕ_H^s , i.e., $\phi_H^s(\text{graph}(c + v_x)) \subset \text{graph}(c + v_x)$ for all $s \leq 0$.*

2.2 α -limit point

Let γ^* be a one-sided global minimizing curve. Consider α -limit points of γ^* , where $x^* \in \mathbb{T}^n$ is called an α -limit point if there exists a monotone sequence $\tau_j \rightarrow -\infty$ as $j \rightarrow \infty$ for which $\gamma^*(\tau_j) \rightarrow x^*$ as $j \rightarrow \infty$. Fix c and define the set for each viscosity solution v of (1.1) as

$$\mathcal{M}_\alpha(v; c) := \overline{\{x^* \in \mathbb{T}^n \mid x^*: \alpha\text{-limit point of } \gamma^*, \gamma^*: \text{minimizer for } v(x), x \in \mathbb{T}^n\}},$$

and take their union,

$$\mathcal{M}_\alpha(c) := \bigcup_v \overline{\mathcal{M}_\alpha(v; c)},$$

where \overline{A} stands for the closure of $A \subset \mathbb{R}^m$ for $m \in \mathbb{N}$, and the union is taken over all the viscosity solutions of (1.1) with fixed c . Here are properties of $\mathcal{M}_\alpha(v; c)$:

- Theorem 2.2.** (i) *Each viscosity solution v to (1.1) is differentiable on $\mathcal{M}_\alpha(v; c)$.*
- (ii) *Let γ^* be a minimizing curve of v and (x^*, ξ^*) be an α -limit point of $(\gamma^*(s), \gamma^{*'}(s))$. Then $\xi^* = H_p(x^*, c + v_x(x^*))$.*
- (iii) *Let x^* be a point of $\mathcal{M}_\alpha(v; c)$ and $x(s)$ be the solution of the Euler-Lagrange equation (2.2) with $x(0) = x^*$, $x'(0) = H_p(x^*, c + v_x(x^*))$. Then $x(s) \in \mathcal{M}_\alpha(v; c)$ and $v_x(x(s)) = L_\xi(x(s), x'(s)) - c$ for all $s \in \mathbb{R}$.*
- (iv) *Let v and \tilde{v} be two viscosity solutions of (1.1). If $v = \tilde{v}$ on $\mathcal{M}_\alpha(c)$, then $v = \tilde{v}$ on \mathbb{T}^n .*

Proof. Let x^* be an arbitrary point of $\mathcal{M}_\alpha(v; c)$ such that there exist a minimizing curve $\gamma^* : (-\infty, 0] \rightarrow \mathbb{T}^n$ and $\tau_j \rightarrow -\infty$ ($j \rightarrow \infty$) satisfying $\gamma^*(\tau_j) \rightarrow x^*$ ($j \rightarrow \infty$). Since $\gamma^{*'}(\tau_j) = H_p(\gamma^*(\tau_j), c + v_x(\gamma^*(\tau_j)))$ is a bounded sequence, there exists a subsequence, still denoted by $\gamma^{*'}(\tau_j)$, which converges to some $\xi^* \in \mathbb{R}^n$. Let $x(s) : \mathbb{R} \rightarrow \mathbb{T}^n$ be the

solution of (2.2) with the initial condition $x(0) = x^*$, $x'(0) = \xi^*$. It follows from the continuous dependence on initial conditions that for any $\alpha > 0$,

$$(\gamma^*(\cdot + \tau_j), \gamma^{*'}(\cdot + \tau_j)) \rightarrow (x(\cdot), x'(\cdot)) \text{ uniformly on } [-\alpha, \alpha] \text{ (} j \rightarrow \infty \text{)}.$$

Hence, we see that $x(s) \in \mathcal{M}_\alpha(v; c)$ for all $s \in \mathbb{R}$ by the definition of $\mathcal{M}_\alpha(v; c)$. Since $\gamma^*(\cdot + \alpha + \tau_j) : [-2\alpha, 0] \rightarrow \mathbb{T}^n$ is the unique minimizing curve for $v(\gamma^*(\alpha + \tau_j))$ for each fixed large j , i.e.,

$$\begin{aligned} v(\gamma^*(\alpha + \tau_j)) &= \int_{-2\alpha}^0 (L(\gamma^*(s + \alpha + \tau_j), \gamma^{*'}(s + \alpha + \tau_j)) - c \cdot \gamma^{*'}(s + \alpha + \tau_j) \\ &\quad + h(c)) ds + v(\gamma^*(-\alpha + \tau_j)), \end{aligned}$$

we obtain by letting $j \rightarrow \infty$,

$$v(x(\alpha)) = \int_{-2\alpha}^0 (L(x(s + \alpha), x'(s + \alpha)) - c \cdot x'(s + \alpha) + h(c)) ds + v(x(-\alpha)).$$

Therefore, we see that $x(s + \alpha)$, $s \in [-\alpha, \alpha]$ is a minimizing curve for $v(x(\alpha))$ and thus $v_x(x(\alpha + s)) = L_\xi(x(\alpha + s), x'(\alpha + s)) - c$ for all $s \in [-2\alpha, 0]$ due to (2.3). For $s = -\alpha$, we obtain $v_x(x^*) = L_\xi(x^*, \xi^*) - c$ and $\xi^* = x'(0) = H_p(x^*, c + v_x(x^*))$.

If x^* is an accumulating point of $\mathcal{M}_\alpha(v; c)$, we have a sequence $\{(x_i^*, \xi_i^*)\}$ of α -limit points of minimizing curves $(\gamma_i^*(s), \gamma_i^{*'}(s))$ such that $(x_i^*, \xi_i^*) \rightarrow (x^*, \xi^*)$ as $i \rightarrow \infty$. Let $x(s)$ be the solution of (2.2) with the initial condition (x^*, ξ^*) . Since the above argument holds for each (x_i^*, ξ_i^*) , the continuous dependence yields the same result for $x(s)$. Noting that $\alpha > 0$ is arbitrary, we complete the proof of (i) to (iii).

We prove (iv). Let x be an arbitrary point of \mathbb{T}^n . Let γ^* (resp., $\tilde{\gamma}^*$) be a minimizing curve for $v(x)$ (resp., $\tilde{v}(x)$). There exists $\tau_j \rightarrow -\infty$ ($j \rightarrow \infty$) and $x^* \in \mathcal{M}_\alpha(c)$ such that $\gamma^*(\tau_j) \rightarrow x^*$ ($j \rightarrow \infty$) (resp., $\tilde{\tau}_j \rightarrow -\infty$ ($j \rightarrow \infty$) and $\tilde{x}^* \in \mathcal{M}_\alpha(c)$ such that $\tilde{\gamma}^*(\tilde{\tau}_j) \rightarrow \tilde{x}^*$ ($j \rightarrow \infty$)). It follows from the variational representation formula (2.1) that for each j ,

$$\tilde{v}(x) - v(x) \leq \tilde{v}(\gamma^*(\tau_j)) - v(\gamma^*(\tau_j)) \quad (\text{resp., } \tilde{v}(x) - v(x) \geq \tilde{v}(\tilde{\gamma}^*(\tilde{\tau}_j)) - v(\tilde{\gamma}^*(\tilde{\tau}_j))).$$

Since $v(x^*) = \tilde{v}(x^*)$ (resp., $v(\tilde{x}^*) = \tilde{v}(\tilde{x}^*)$), we conclude $v(x) \leq \tilde{v}(x)$ (resp., $v(x) \geq \tilde{v}(x)$) by letting $j \rightarrow \infty$. \square

By Theorem 2.2, the following sets are well-defined:

$$\begin{aligned} \tilde{\mathcal{M}}_\alpha(v; c) &:= \{(x, H_p(c + v_x(x))) \mid x \in \mathcal{M}_\alpha(v; c)\}, & \tilde{\mathcal{M}}_\alpha(c) &:= \bigcup_v \tilde{\mathcal{M}}_\alpha(v; c), \\ \tilde{\mathcal{M}}_\alpha^*(v; c) &:= \{(x, c + v_x(x)) \mid x \in \mathcal{M}_\alpha(v; c)\}, & \tilde{\mathcal{M}}_\alpha^*(c) &:= \bigcup_v \tilde{\mathcal{M}}_\alpha^*(v; c), \end{aligned}$$

where the union in the above is taken over all the viscosity solutions of (1.1) with fixed c .

Remark 2.3. The curve $x(s)$ in Theorem 2.2 is a global minimizing curve of v , which is not a homoclinic/heteroclinic orbit. The set $\mathcal{M}_\alpha(v; c)$ is equal to the closure of the family of all such global minimizing curves of v , and

$$\begin{aligned} \tilde{\mathcal{M}}_\alpha(v; c) &= \overline{\{(x^*, \xi^*) \mid (x^*, \xi^*): \alpha\text{-limit point of } (\gamma^*(s), \gamma^{*'}(s)), \\ &\quad \gamma^*: \text{minimizing curve of } v\}}. \end{aligned}$$

Here are properties of $\tilde{\mathcal{M}}_\alpha(v; c)$ and $\tilde{\mathcal{M}}_\alpha^*(v; c)$:

Theorem 2.4. (i) For each point $x \in \mathbb{T}^n$, there exists $\xi \in \mathbb{R}^n$ such that $\phi_L^s(x, \xi)$ falls into $\tilde{\mathcal{M}}_\alpha(v; c)$ as $s \rightarrow -\infty$, i.e., any accumulating point of $\{\phi_L^s(x, \xi)\}_{s \leq 0}$ belongs to $\tilde{\mathcal{M}}_\alpha(v; c)$. If $v_x(x)$ exists, $\xi = H_p(x, c + v_x(x))$.

(ii) $\tilde{\mathcal{M}}_\alpha(v; c)$ is a ϕ_L^s -invariant subset of $\mathbb{T}^n \times \mathbb{R}^n$, i.e., $\phi_L^s(\tilde{\mathcal{M}}_\alpha(v; c)) = \tilde{\mathcal{M}}_\alpha(v; c)$ for all $s \in \mathbb{R}$.

(iii) For each point $x \in \mathbb{T}^n$, there exists $p \in \mathbb{R}^n$ such that $\phi_H^s(x, p)$ falls into $\tilde{\mathcal{M}}_\alpha^*(v; c)$ as $s \rightarrow -\infty$ along $\text{graph}(c + v_x)$. If $v_x^\varepsilon(x)$ exists, $p = c + v_x^\varepsilon(x)$.

(iv) $\tilde{\mathcal{M}}_\alpha^*(v; c)$ is a ϕ_H^s -invariant subset of $\text{graph}(c + v_x)$.

Proof. (i) is clear by definition. (ii) is clear by (iii) of Theorem 2.2. Set $p := L_\xi(x, \xi)$, $\xi := \gamma^{*'}(0)$ with a minimizing curve γ^* for $v(x)$. Then we have $\phi_H^s(x, p) = (\pi \circ \phi_L^s(x, \xi), L_\xi(\phi_L^s(x, \xi)))$, where $\pi : \mathbb{T}^n \times \mathbb{R}^n \rightarrow \mathbb{T}^n$ is the standard projection. Hence, (iii) follows from above (i), (2.3) with equivalence between (2.2) and (2.4), and (ii) of Theorem 2.2. Let (x^*, p^*) be an arbitrary point of $\tilde{\mathcal{M}}_\alpha^*(v; c)$. (iv) follows from (iii) of Theorem 2.2 with $x'(0) = H_p(x^*, p^*)$ and equivalence between (2.2) and (2.4). \square

2.3 Comparison with Mather set and Aubry set

In this subsection, we compare $\mathcal{M}_\alpha(c)$, $\tilde{\mathcal{M}}_\alpha(c)$ with the (projected) Mather set and the (projected) Aubry set defined in weak KAM theory.

We recall the measure-theoretical aspect of weak KAM theory, which is first investigated in [21]. Let $\gamma^* : (-\infty, 0] \rightarrow \mathbb{T}^n$ be a minimizing curve for $v(x)$. Note that $(\gamma^*(s), \gamma^{*'}(s))$ is contained in a compact set $K \subset \mathbb{T}^n \times \mathbb{R}^n$ for each $s \leq 0$, because $\{v^\varepsilon\}_{\varepsilon > 0}$ is equi-Lipschitz continuous. Define the linear functional for each $T > 0$

$$(2.5) \quad \Psi_T(\varphi) := \frac{1}{T} \int_{-T}^0 \varphi(\gamma^*(s), \gamma^{*'}(s)) ds, \quad \varphi \in C_c(\mathbb{T}^n \times \mathbb{R}^n, \mathbb{R}),$$

where $C_c(\mathbb{T}^n \times \mathbb{R}^n, \mathbb{R})$ denotes the family of compactly supported continuous functions. Then, by the Riesz representation theorem, there exists a probability measure μ_T on $\mathbb{T}^n \times \mathbb{R}^n$ such that

$$(2.6) \quad \Psi_T(\varphi) = \int_{\mathbb{T}^n \times \mathbb{R}^n} \varphi(x, \xi) d\mu_T \quad \text{for all } \varphi \in C_c(\mathbb{T}^n \times \mathbb{R}^n, \mathbb{R}).$$

Since the support of μ_T , denoted by $\text{supp}(\mu_T)$, is contained in K for all $T > 0$, we have $T_i \rightarrow \infty$ as $i \rightarrow \infty$ for which μ_{T_i} converges weakly to a probability measure μ^* . We see that μ^* is ϕ_L^s -invariant. Furthermore, (2.6) with $\varphi(x, \xi) = L(x, \xi) - c \cdot \xi + h(c)$ (more precisely, the right-hand-side is re-defined to be 0 continuously outside K so that it belongs to $C_c(\mathbb{T}^n \times \mathbb{R}^n, \mathbb{R})$, which does not change anything in regards to a measure supported in K) yields

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, \xi) - c \cdot \xi + h(c) d\mu_T = \frac{1}{T} (v(\gamma^*(0)) - v(\gamma^*(-T))).$$

Putting $T = T_i$ and sending $i \rightarrow \infty$, we obtain

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, \xi) - c \cdot \xi \, d\mu^* = -h(c).$$

We observe the minimizing property of μ^* . For any $(x, \xi) \in \mathbb{T}^n \times \mathbb{R}^n$ and $t > \tau > 0$, we have

$$\begin{aligned} (2.7) \quad v(\pi \circ \phi_L^t(x, \xi)) &\leq \int_{-\tau}^t L^c(\phi_L^s(x, \xi)) + h(c) \, ds + v(\pi \circ \phi_L^{-\tau}(x, \xi)) \\ &= \int_{-\tau-t}^0 L^c(\phi_L^{s+t}(x, \xi)) + h(c) \, ds + v(\pi \circ \phi_L^{-\tau}(x, \xi)). \end{aligned}$$

Then, integrating the inequality with any ϕ_L^s -invariant probability measure μ defined on $\mathbb{T}^n \times \mathbb{R}^n$, we have

$$0 \leq \int_{\mathbb{T}^n \times \mathbb{R}^n} L^c(x, \xi) + h(c) \, d\mu \Leftrightarrow \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, \xi) - c \cdot \xi \, d\mu \geq -h(c).$$

Therefore we conclude that μ^* is a minimizing measure for the minimizing problem (M) in Introduction.

Taking $\phi(x, \xi) := \psi_x(x) \cdot \xi$ (in K , otherwise re-defined to be 0 continuously) with $\psi \in C^1(\mathbb{T}^n)$ in (2.6), we have

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \psi_x(x) \cdot \xi \, d\mu_T = \frac{1}{T} \int_{-T}^0 \psi_x(\gamma^*(s)) \cdot \gamma^{*'}(s) \, ds = \frac{1}{T} \{\psi(\gamma^*(0)) - \psi(\gamma^*(-T))\}.$$

With $T = T_j$ and $j \rightarrow \infty$, we see that $\mu = \mu^*$ satisfies

$$(2.8) \quad \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi_x(x) \cdot \xi \, d\mu = 0 \quad \text{for all } \psi \in C^1(\mathbb{T}^n).$$

A probability measure μ satisfying (2.8) is said to be *holonomic*. It is proved that the minimizing problem (1.3) within all ϕ_L^s -invariant measures is equivalent to the one within all holonomic measures [20].

The (projected) Mather set $\mathcal{M}(c) := \pi \tilde{\mathcal{M}}(c)$, $\tilde{\mathcal{M}}(c)$ is defined as

$$\tilde{\mathcal{M}}(c) := \overline{\bigcup_{\mu} \text{supp}(\mu)},$$

where the union is taken over all minimizing measures of (M). As shown in [13], [11], for each point $(x, \xi) \in \tilde{\mathcal{M}}(c)$, any viscosity solution v of (1.1) is differentiable at x possessing the common derivative

$$(2.9) \quad v_x(x) = L_\xi(x, \xi) - c.$$

In fact, since (2.7) holds for each $(x, \xi) \in \mathbb{T}^n \times \mathbb{R}^n$, it must be an equality on the support of each minimizing measure μ , because otherwise our integration with μ yields $0 < \int_{\mathbb{T}^n \times \mathbb{R}^n} L^c(x, \xi) \, d\mu + h(c)$. Continuity implies that (2.7) is an equality for each $(x, \xi) \in \tilde{\mathcal{M}}(c)$. Therefore $\pi \circ \phi_L^{+t}(x, \xi) : [-\tau - t, 0] \rightarrow \mathbb{T}^n$ yields a minimizing curve for $v(\phi_L^t(x, \xi))$ and v is differentiable at $\pi \circ \phi_L^0(x, \xi) = (x, \xi)$ with (2.9).

The (projected) Aubry set $\mathcal{A}(c) := \pi \tilde{\mathcal{A}}(c)$, $\tilde{\mathcal{A}}(c)$ is defined through conjugate pairs of weak KAM solutions (see [13] for detail). By definition, for each point $(x, \xi) \in \tilde{\mathcal{A}}(c)$, any viscosity solution v of (1.1) is differentiable at x possessing the common derivative (2.9). The following characterization of the projected Aubry set [13] is useful:

$$\mathcal{A}(c) = \{x \in \mathbb{T}^n \mid \mathfrak{h}^c(x, x) = 0\},$$

$$\mathfrak{h}^c(x, y) := \liminf_{T \rightarrow \infty} \inf_{\gamma \in AC, \gamma(0)=x, \gamma(-T)=y} \int_{-T}^0 L(\gamma(s), \gamma'(s)) - c \cdot \gamma'(s) + h(c) ds,$$

where \mathfrak{h}^c is the Peierls barrier.

Theorem 2.5. (i) *Let μ be any minimizing measure of (M). Then, every point $(x, \xi) \in \text{supp}(\mu)$ itself is an α -limit point of $\phi_L^s(x, \xi)$.*

$$(ii) \quad \mathcal{M}(c) \subset \mathcal{M}_\alpha(c) \subset \mathbb{T}^n, \quad \tilde{\mathcal{M}}(c) \subset \tilde{\mathcal{M}}_\alpha(c) \subset \mathbb{T}^n \times \mathbb{R}^n.$$

$$(iii) \quad \mathcal{M}_\alpha(c) \subset \mathcal{A}(c) \subset \mathbb{T}^n, \quad \tilde{\mathcal{M}}_\alpha(c) \subset \tilde{\mathcal{A}}(c) \subset \mathbb{T}^n \times \mathbb{R}^n.$$

Proof. It follows from Poincaré's recurrence theorem that μ -a.e. points of $\text{supp}(\mu)$ are recurrent, which yields (ii). A slightly more detailed argument is necessary to prove that every point is recurrent. We show a direct proof of (i) and (ii).

Let μ be any minimizing measure of (M). As we already observed, for any $(x, \xi) \in \mathbb{T}^n \times \mathbb{R}^n$, we have (2.7), which must be an equality on $\text{supp}(\mu)$. Hence, for each $(x, \xi) \in \text{supp}(\mu)$, the curve $\pi \circ \phi_L^s(x, \xi) : [-T, 0] \rightarrow \mathbb{T}^n$ is the unique minimizing curve for $v(x)$.

Suppose that there exists a minimizing measure μ of (M) for which we have a point $(x^*, \xi^*) \in \text{supp}(\mu) \setminus \tilde{\mathcal{M}}_\alpha(c)$. Then, (x^*, ξ^*) cannot be any α -limit point of $\phi_L^s(x^*, \xi^*)$. Therefore, we have a closed δ -ball B_δ^* of (x^*, ξ^*) to which $\phi_L^s(x^*, \xi^*)$ never comes back for $s \rightarrow -\infty$. The following two cases are possible:

- (a) Taking smaller $\delta > 0$ if necessary, the set $B_\delta^* \setminus \{\phi_L^s(x^*, \xi^*)\}_{s \in \mathbb{R}}$ contains no other points of $\text{supp}(\mu)$.
- (b) There exists a sequence $(x_i, \xi_i) \in [B_\delta^* \cap \text{supp}(\mu)] \setminus \{\phi_L^s(x^*, \xi^*)\}_{s \in \mathbb{R}}$ such that $(x_i, \xi_i) \rightarrow (x^*, \xi^*)$ as $i \rightarrow \infty$.

In case (a), taking a continuous function $\varphi : \mathbb{T}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ such that $\text{supp}(\varphi) \subset B_\delta^*$ and $\varphi > 0$ in $\text{supp}(\varphi)$, we have for each $T > 0$,

$$(2.10) \quad \int_{\mathbb{T}^n \times \mathbb{R}^n} \left\{ \frac{1}{T} \int_{-T}^0 \varphi(\phi_L^s(x, \xi)) ds \right\} d\mu = \frac{1}{T} \int_{-T}^0 \left\{ \int_{\mathbb{T}^n \times \mathbb{R}^n} \varphi(\phi_L^s(x, \xi)) d\mu \right\} ds$$

$$= \frac{1}{T} \int_{-T}^0 \left\{ \int_{\mathbb{T}^n \times \mathbb{R}^n} \varphi(\phi_L^0(x, \xi)) d\mu \right\} ds = \int_{\mathbb{T}^n \times \mathbb{R}^n} \varphi(x, \xi) d\mu > 0.$$

Since $\phi_L^s(x, \xi)$ with $(x, \xi) \in \text{supp}(\mu) \cap B_\delta^*$ never comes back to B_δ^* for $s \rightarrow -\infty$, we have $\frac{1}{T} \int_{-T}^0 \varphi(\phi_L^s(x, \xi)) ds \rightarrow 0$ as $T \rightarrow \infty$ pointwise on $\text{supp}(\mu) \cap B_\delta^*$ and hence, the left-hand-side of (2.10) is to be 0, which is a contradiction.

In case (b), we cannot find any subsequence of (x_i, ξ_i) , still denoted by (x_i, ξ_i) , such that $(x_i, \xi_i) \in \tilde{\mathcal{M}}_\alpha(c)$, because otherwise (x^*, ξ^*) must belong to $\tilde{\mathcal{M}}_\alpha(c)$. Hence, taking

smaller $\delta > 0$ if necessary, all points of $[B_\delta^* \cap \text{supp}(\mu)] \setminus \{\phi_L^s(x^*, \xi^*)\}_{s \in \mathbb{R}}$ do not belong to $\tilde{\mathcal{M}}_\alpha(c)$. Note that if $(x, \xi) \in \text{supp}(\mu)$, the α -limit points of $\phi_L^s(x, \xi)$ belong to $\text{supp}(\mu)$. Therefore, for each $(x, \xi) \in [B_\delta^* \cap \text{supp}(\mu)] \setminus \{\phi_L^s(x^*, \xi^*)\}_{s \in \mathbb{R}}$, the α -limit points of $\phi_L^s(x, \xi)$ cannot belong to $[B_\delta^* \cap \text{supp}(\mu)] \setminus \{\phi_L^s(x^*, \xi^*)\}_{s \in \mathbb{R}}$, nor to $B_\delta^* \cap \{\phi_L^s(x^*, \xi^*)\}_{s \in \mathbb{R}}$ because otherwise $\{\phi_L^s(x^*, \xi^*)\}_{s \in \mathbb{R}} \subset \mathcal{M}_\alpha(c)$ due to (iii) of Theorem 2.2. Thus, since $\phi_L^s(x, \xi)$ with $(x, \xi) \in \text{supp}(\mu) \cap B_\delta^*$ never comes back to B_δ^* for $s \rightarrow -\infty$, the same argument with (2.10) as that in the case of (a) yields a contradiction.

We prove (iii). It is enough to show $\mathcal{M}_\alpha(c) \subset \mathcal{A}(c)$, because if $x \in \mathcal{M}_\alpha(c) \cap \mathcal{A}(c)$, we have a viscosity solution v for which $(x, \xi) = (x, H_p(x, c + v_x(x))) \in \tilde{\mathcal{M}}_\alpha(v; c)$ and hence $(x, \xi) \in \tilde{\mathcal{A}}(c)$ due to (2.9) on $\tilde{\mathcal{A}}(c)$ for any viscosity solution. Let x be an arbitrary point of $\mathcal{M}_\alpha(c)$ for which we have a minimizing curve $\gamma : (-\infty, 0] \rightarrow \mathbb{T}^n$ and $T_i \rightarrow \infty$ such that $\gamma(-T_i) \rightarrow x$ as $i \rightarrow \infty$. Then, it holds that

$$v(\gamma(-T_i)) = \int_{-T_{i+j}}^{-T_i} L(\gamma(s), \gamma'(s)) - c \cdot \gamma'(s) + h(c) ds + v(\gamma(-T_{i+j}))$$

or, with $\gamma_i(s) := \gamma(s - T_i) : [-T_{i+j} + T_i, 0] \rightarrow \mathbb{T}^n$,

$$v(\gamma_i(0)) = \int_{-T_{i+j}+T_i}^0 L(\gamma_i(s), \gamma_i'(s)) - c \cdot \gamma_i'(s) + h(c) ds + v(\gamma_i(-T_{i+j} + T_i)).$$

Define

$$\gamma_{ij}(s) := \begin{cases} \gamma_i(s) & \text{for } s \in [-T_{i+j} + T_i, 0], \\ \gamma_i(-T_{i+j} + T_i) + \frac{\gamma_i(-T_{i+j} + T_i) - x}{\varepsilon_{ij}} \{s - (-T_{i+j} + T_i)\} & \text{for } s \in [-T_{i+j} + T_i - \varepsilon_{ij}, -T_{i+j} + T_i], \end{cases}$$

where $\varepsilon_{ij} := \sqrt{|\gamma_i(-T_{i+j} + T_i) - x|}$. Then we see that

$$\begin{aligned} \mathfrak{h}^c(\gamma_i(0), x) &\leq \liminf_{j \rightarrow \infty} \int_{-T_{i+j}+T_i-\varepsilon_{ij}}^0 L(\gamma_{ij}(s), \gamma_{ij}'(s)) - c \cdot \gamma_{ij}'(s) + h(c) ds \\ &= \liminf_{j \rightarrow \infty} v(\gamma_i(0)) - v(\gamma_i(-T_{i+j} + T_i)) + O(\varepsilon_{ij}) = v(\gamma(-T_i)) - v(x). \end{aligned}$$

Since $\gamma(-T_i) \rightarrow x$ as $i \rightarrow \infty$ and continuity of \mathfrak{h}^c , we have $\mathfrak{h}^c(x, x) \leq 0$. On the other hand, for a minimizing curve γ^* of the variational problem

$$\inf_{\gamma \in AC, \gamma(0)=x, \gamma(-T)=x} \int_{-T}^0 L(\gamma(s), \gamma'(s)) - c \cdot \gamma'(s) + h(c) ds,$$

we have

$$\int_{-T}^0 L(\gamma^*(s), \gamma^{*'}(s)) - c \cdot \gamma^{*'}(s) + h(c) ds \geq v(\gamma^*(0)) - v(\gamma^*(-T)) = 0$$

for each $T > 0$, which implies $\mathfrak{h}^c(x, x) \geq 0$. Hence, we obtain $\mathfrak{h}^c(x, x) = 0$ and therefore $x \in \mathcal{A}(c)$. Continuity of \mathfrak{h}^c implies $\mathfrak{h}^c(x, x) = 0$ also for accumulating points of $\mathcal{M}_\alpha(c)$. Thus, we conclude $\mathcal{M}_\alpha(c) \subset \mathcal{A}(c)$, $\tilde{\mathcal{M}}_\alpha(c) \subset \tilde{\mathcal{A}}(c)$. \square

Since any viscosity solution of (1.1) is differentiable on the projected Aubry set, we have

Corollary 2.6. *Any viscosity solution v of (1.1) is differentiable on $\mathcal{M}_\alpha(c)$ with the common derivative $v_x(x) = L_\xi(x, \xi) - c$ for each point $(x, \xi) \in \tilde{\mathcal{M}}_\alpha(c)$.*

This is apparently not trivial from the viscosity-solution-wise definition of $\mathcal{M}_\alpha(c)$, $\tilde{\mathcal{M}}_\alpha(c)$.

Remark 2.7. (i) $\mathcal{M}_\alpha(c)$ can be strictly smaller than $\mathcal{A}(c)$ in general.

(ii) Currently it is not clear if $\mathcal{M}(c)$ is strictly smaller than $\mathcal{M}_\alpha(c)$ in general or not.

In fact, (i) of Remark 2.7 is the case, when there are homoclinic/heteroclinic orbits. For instance, suppose that Hamiltonian dynamics with $n = 1$ has a single stationary solution x_0 with a homoclinic orbit γ^* moving over $\mathbb{T} \setminus \{x_0\}$. There exists a value c such that $\text{graph}(c + v_x) = \{(\gamma^*(s), c + v_x(\gamma^*(s)))\}_{s \in \mathbb{R}}$. Then, $\mathcal{A}(c)$ is equal to $\{\gamma^*(s)\}_{s \in \mathbb{R}} \cup \{x_0\} = \mathbb{T}$, and $\mathcal{M}_\alpha(c)$ is equal to $\{x_0\}$.

In regards to (ii) of Remark 2.7, one could expect the case where $\mathcal{M}(c)$ is strictly smaller than $\mathcal{M}_\alpha(c)$ because of the following observation. Let $\gamma^*(s)$ be a minimizing curve and (x^*, ξ^*) be an α -limit point of $(\gamma^*(s), \gamma^{*'}(s))$. Let $\tau(T)$ denote the total length of time for which $(\gamma^*(s), \gamma^{*'}(s))$ stays in a δ -ball of (x^*, ξ^*) within $s \in [-T, 0]$. It seems that $r(\delta) := \lim_{T \rightarrow \infty} \tau(T)/T$ determines whether or not (x^*, ξ^*) is a point of the support of the minimizing measure induced by (2.5) and (2.6), namely, (x^*, ξ^*) would not belong to the support, if $r(\delta) = 0$ for some $\delta > 0$.

3 Weak KAM theory for discount problem

We extend what we observed in Section 2 to the discount problem. The family of α -limit points of minimizing curves has properties similar to those of $\mathcal{M}_\alpha(c)$, $\tilde{\mathcal{M}}_\alpha(c)$. We present certain minimizing measures to relate them to the family of α -limit points. These minimizing measures are different from the ones used in [7], [22] for the selection criterion in the vanishing discount process.

3.1 Viscosity solution and minimizing curve

Let v^ε be the unique viscosity solution of (1.2). Note that v^ε is Lipschitz continuous and semiconcave. It is well-known that v^ε satisfies for each $x \in \mathbb{T}^n$ and $T > 0$,

$$(3.1) \quad v^\varepsilon(x) = \inf_{\gamma \in AC, \gamma(0)=x} \left\{ \int_{-T}^0 e^{\varepsilon s} (L(\gamma(s), \gamma'(s)) - c \cdot \gamma'(s) + h(c)) ds + e^{-\varepsilon T} v^\varepsilon(\gamma(-T)) \right\}.$$

It follows from Tonelli's theory that there exists a minimizing curve $\gamma^\varepsilon : [-T, 0] \rightarrow \mathbb{T}^n$ for $v^\varepsilon(x)$, which is a C^2 -solution of the Euler-Lagrange equation generated by $L^{c, \varepsilon} := e^{\varepsilon s} \{L - c \cdot \xi + h(c)\}$,

$$(3.2) \quad \begin{aligned} \frac{d}{ds} \{L_\xi^{c, \varepsilon}(x(s), x'(s))\} &= L_x^{c, \varepsilon}(x(s), x'(s)) \\ \Leftrightarrow \frac{d}{ds} \{L_\xi(x(s), x'(s))\} &= L_x(x(s), x'(s)) - \varepsilon L_\xi(x(s), x'(s)) + \varepsilon c. \end{aligned}$$

The flow of (3.2) is denoted by $\phi_{L,c,\varepsilon}^s$, i.e., $\phi_{L,c,\varepsilon}^s(x(0), x'(0)) = (x(s), x'(s))$. Set $p^\varepsilon(s) := L_\xi(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s))$. Then $(\gamma^\varepsilon(s), p^\varepsilon(s))$ solves the equation,

$$(3.3) \quad \begin{cases} x'(s) = H_p(x(s), p(s)), \\ p'(s) = -H_x(x(s), p(s)) + \varepsilon c - \varepsilon p(s). \end{cases}$$

We see that (3.2), (3.3) can be regarded as (2.2), (2.4) with a friction term. The flow of (3.3) is denoted by $\phi_{H,c,\varepsilon}^s$.

Proposition 3.1. *Let $\gamma^\varepsilon : [-T, 0] \rightarrow \mathbb{T}^n$ be a minimizing curve for $v^\varepsilon(x)$. For each $\tau \in [0, T)$, we have*

$$\begin{aligned} & v^\varepsilon(\gamma^\varepsilon(-\tau)) \\ &= e^{\varepsilon\tau} \left\{ \int_{-T}^{-\tau} e^{\varepsilon s} (L(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) - c \cdot \gamma^{\varepsilon'}(s) + h(c)) ds + e^{-\varepsilon T} v^\varepsilon(\gamma^\varepsilon(-T)) \right\} \\ &= \int_{-T+\tau}^0 e^{\varepsilon s} (L(\gamma^\varepsilon(s-\tau), \gamma^{\varepsilon'}(s-\tau)) - c \cdot \gamma^{\varepsilon'}(s-\tau) + h(c)) ds + e^{\varepsilon(-T+\tau)} v^\varepsilon(\gamma^\varepsilon(-T)). \end{aligned}$$

This is well-known as a dynamic programming principle for the discount value function (3.1). It follows from Proposition 3.1 that we have

$$(3.4) \quad \begin{aligned} v^\varepsilon(x) &= \int_{-\tau}^0 e^{\varepsilon s} (L(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) - c \cdot \gamma^{\varepsilon'}(s) + h(c)) ds \\ &\quad + e^{-\varepsilon\tau} v^\varepsilon(\gamma^\varepsilon(-\tau)) \quad \text{for each } \tau \in [0, T]. \end{aligned}$$

Proposition 3.2. *Let $\gamma^\varepsilon : [-T, 0] \rightarrow \mathbb{T}^n$ be a minimizing curve for $v^\varepsilon(x)$. Then v^ε is differentiable on $\gamma^\varepsilon(s)$ for all $s \in [-T, 0)$ satisfying*

$$v_x^\varepsilon(\gamma^\varepsilon(s)) = L_\xi(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) - c.$$

If $v_x^\varepsilon(x)$ exists, this holds for $s = 0$ and γ^ε is the unique minimizing curve for $v^\varepsilon(x)$.

Proof. Since v^ε is semiconcave, the superdifferential is a non-empty set: $D^+v^\varepsilon(\gamma^\varepsilon(s)) \neq \emptyset$ for all $s \in [-T, 0]$ (see, e.g., [6]). Hence it is enough to check that the subdifferential contains $L_\xi(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) - c$. For any $\tau \in [-T, 0)$, we have

$$\begin{aligned} v^\varepsilon(x) &= \int_{-T}^0 e^{\varepsilon s} (L(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) - c \cdot \gamma^{\varepsilon'}(s) + h(c)) ds + e^{-\varepsilon T} v^\varepsilon(\gamma^\varepsilon(-T)) \\ &= \int_{\tau}^0 e^{\varepsilon s} (L(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) - c \cdot \gamma^{\varepsilon'}(s) + h(c)) ds + e^{\varepsilon\tau} v^\varepsilon(\gamma^\varepsilon(\tau)). \end{aligned}$$

Take $\delta > 0$ small so that $\tau + \delta \in (-T, 0)$. Introduce a continuous curve γ with $w \in \mathbb{R}^n$ as

$$\gamma(s) := \begin{cases} \gamma^\varepsilon(s) & \text{for } \tau + \delta \leq s \leq 0, \\ \gamma^\varepsilon(s) + (\tau + \delta - s) \frac{w}{\delta} & \text{for } \tau \leq s \leq \tau + \delta. \end{cases}$$

Due to the variational property, we have

$$v^\varepsilon(x) \leq \int_{\tau}^0 e^{\varepsilon s} (L(\gamma(s), \gamma'(s)) - c \cdot \gamma'(s) + h(c)) ds + e^{\varepsilon\tau} v^\varepsilon(\gamma(\tau)).$$

Hence, we obtain

$$\begin{aligned}
0 &\leq \int_{\tau}^{\tau+\delta} e^{\varepsilon s} \left[\left\{ L(\gamma^{\varepsilon}(s) + (\tau + \delta - s) \frac{w}{\delta}, \gamma^{\varepsilon'}(s) - \frac{w}{\delta}) - c \cdot (\gamma^{\varepsilon'}(s) - \frac{w}{\delta}) \right\} \right. \\
&\quad \left. - \left\{ L(\gamma^{\varepsilon}(s), \gamma^{\varepsilon'}(s)) - c \cdot \gamma^{\varepsilon'}(s) \right\} \right] ds + e^{\varepsilon \tau} \{ v^{\varepsilon}(\gamma^{\varepsilon}(\tau) + w) - v^{\varepsilon}(\gamma^{\varepsilon}(\tau)) \} \\
&= e^{\varepsilon(\tau+\theta\delta)} \left[\left\{ L(\gamma^{\varepsilon}(\tau + \theta\delta) + (\delta - \theta\delta) \frac{w}{\delta}, \gamma^{\varepsilon'}(\tau + \theta\delta) - \frac{w}{\delta}) - c \cdot (\gamma^{\varepsilon'}(\tau + \theta\delta) - \frac{w}{\delta}) \right\} \right. \\
&\quad \left. - \left\{ L(\gamma^{\varepsilon}(\tau + \theta\delta), \gamma^{\varepsilon'}(\tau + \theta\delta)) - c \cdot \gamma^{\varepsilon'}(\tau + \theta\delta) \right\} \right] \delta \\
&\quad + e^{\varepsilon \tau} \{ v^{\varepsilon}(\gamma^{\varepsilon}(\tau) + w) - v^{\varepsilon}(\gamma^{\varepsilon}(\tau)) \} \\
&= e^{\varepsilon(\tau+\theta\delta)} \left\{ L_x(\gamma^{\varepsilon}(\tau + \theta\delta) + \tilde{\theta}(1 - \theta)w, \gamma^{\varepsilon'}(\tau + \theta\delta) - \tilde{\theta} \frac{w}{\delta}) \cdot (1 - \theta)w \right. \\
&\quad \left. + L_{\xi}(\gamma^{\varepsilon}(\tau + \theta\delta) + \tilde{\theta}(1 - \theta)w, \gamma^{\varepsilon'}(\tau + \theta\delta) - \tilde{\theta} \frac{w}{\delta}) \cdot (-\frac{w}{\delta}) + c \cdot \frac{w}{\delta} \right\} \delta \\
&\quad + e^{\varepsilon \tau} \{ v^{\varepsilon}(\gamma^{\varepsilon}(\tau) + w) - v^{\varepsilon}(\gamma^{\varepsilon}(\tau)) \},
\end{aligned}$$

for some $\theta, \tilde{\theta} \in (0, 1)$. Therefore, we have

$$\begin{aligned}
&\frac{v^{\varepsilon}(\gamma^{\varepsilon}(\tau) + w) - v^{\varepsilon}(\gamma^{\varepsilon}(\tau)) - \{L_{\xi}(\gamma^{\varepsilon}(\tau), \gamma^{\varepsilon'}(\tau)) - c\} \cdot w}{|w|} \\
&\geq -e^{\varepsilon\theta\delta} L_x(\gamma^{\varepsilon}(\tau + \theta\delta) + \tilde{\theta}(1 - \theta)w, \gamma^{\varepsilon'}(\tau + \theta\delta) - \tilde{\theta} \frac{w}{\delta}) \cdot (1 - \theta) \frac{w}{|w|} \delta \\
&\quad + e^{\varepsilon\theta\delta} L_{\xi}(\gamma^{\varepsilon}(\tau + \theta\delta) + \tilde{\theta}(1 - \theta)w, \gamma^{\varepsilon'}(\tau + \theta\delta) - \tilde{\theta} \frac{w}{\delta}) \cdot \frac{w}{|w|} - L_{\xi}(\gamma^{\varepsilon}(\tau), \gamma^{\varepsilon'}(\tau)) \cdot \frac{w}{|w|} \\
&\quad + (1 - e^{\varepsilon\theta\delta}) c \cdot \frac{w}{|w|}.
\end{aligned}$$

Taking $\delta = \sqrt{|w|}$ and sending w to 0, we conclude that

$$\liminf_{|w| \rightarrow 0} \frac{v^{\varepsilon}(\gamma^{\varepsilon}(\tau) + w) - v^{\varepsilon}(\gamma^{\varepsilon}(\tau)) - \{L_{\xi}(\gamma^{\varepsilon}(\tau), \gamma^{\varepsilon'}(\tau)) - c\} \cdot w}{|w|} \geq 0,$$

which means that the subdifferential $D^-v^{\varepsilon}(\gamma^{\varepsilon}(\tau))$ contains $L_{\xi}(\gamma^{\varepsilon}(\tau), \gamma^{\varepsilon'}(\tau)) - c$ for all $\tau \in [-T, 0]$.

Similar reasoning with the curve

$$\gamma(s) := \begin{cases} \gamma^{\varepsilon}(s) + \frac{s + \delta}{\delta} w & \text{for } -\delta \leq s \leq 0, \\ \gamma^{\varepsilon}(s) & \text{for } -T \leq s \leq -\delta \end{cases}$$

shows that $D^+v^{\varepsilon}(\gamma^{\varepsilon}(0))$ contains $L_{\xi}(\gamma^{\varepsilon}(0), \gamma^{\varepsilon'}(0)) - c$. If $v_x^{\varepsilon}(\gamma^{\varepsilon}(0)) = v_x^{\varepsilon}(x)$ exists, the set $D^+v^{\varepsilon}(\gamma^{\varepsilon}(0))$ must be singleton and therefore $v_x^{\varepsilon}(\gamma^{\varepsilon}(0)) = L_{\xi}(\gamma^{\varepsilon}(0), \gamma^{\varepsilon'}(0)) - c$. Uniqueness of the discount Euler-Lagrange equation (3.2) yields uniqueness of the minimizing curve. \square

Each minimizing curve $\gamma^{\varepsilon} : [-T, 0] \rightarrow \mathbb{T}^n$ for $v^{\varepsilon}(x)$ can be extended to $(-\infty, 0]$. We call an extended minimizing curve defined on $(-\infty, 0]$ a one-sided global minimizing curve or just minimizing curve. Since $v_x^{\varepsilon}(\gamma^{\varepsilon}(-T)) = L_{\xi}(\gamma^{\varepsilon}(-T), \gamma^{\varepsilon'}(-T)) - c$ due to Proposition 3.2, the minimizing curve for $v(\gamma^{\varepsilon}(-T))$ is uniquely obtained as $\gamma^{\varepsilon}(-T + \cdot)$

on any interval $[-\tilde{T}, 0]$. Therefore, with (3.4), we conclude that for a one-sided global minimizing curve γ^ε for $v^\varepsilon(x)$ we have

$$(3.5) \quad v^\varepsilon(x) = \int_{-\tau}^0 e^{\varepsilon s} (L(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) - c \cdot \gamma^{\varepsilon'}(s) + h(c)) ds \\ + e^{-\varepsilon \tau} v^\varepsilon(\gamma^\varepsilon(-\tau)) \text{ for any } \tau \geq 0.$$

Equivalence between (3.2) and (3.3), and Proposition 3.2 yield

Theorem 3.3. *The set $\text{graph}(c + v_x^\varepsilon)$ is backward invariant under the discount Hamiltonian flow $\phi_{H,c,\varepsilon}^s$, i.e., $\phi_{H,c,\varepsilon}^s(\text{graph}(c + v_x^\varepsilon)) \subset \text{graph}(c + v_x^\varepsilon)$ for all $s \leq 0$.*

3.2 α -limit point

Now we introduce the family of α -limit points of all one-sided global minimizing curves of v^ε :

$$\mathcal{M}_\alpha^\varepsilon(c) := \overline{\{x^\varepsilon \mid x^\varepsilon: \alpha\text{-limit point of } \gamma^\varepsilon, \gamma^\varepsilon: \text{minimizer for } v^\varepsilon(x), x \in \mathbb{T}^n\}}.$$

Here are properties of $\mathcal{M}_\alpha^\varepsilon(c)$:

- Theorem 3.4.** (i) *The viscosity solution v^ε of (1.2) is differentiable on $\mathcal{M}_\alpha^\varepsilon(c)$.*
- (ii) *Let γ^ε be a minimizing curve of v^ε and $(x^\varepsilon, \xi^\varepsilon)$ be an α -limit point of $(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s))$. Then $\xi^\varepsilon = H_p(x^\varepsilon, c + v_x^\varepsilon(x^\varepsilon))$.*
- (iii) *Let x^ε be a point of $\mathcal{M}_\alpha^\varepsilon(c)$ and $x(s)$ be the solution of the discount Euler-Lagrange equation (3.2) with $x(0) = x^\varepsilon$, $x'(0) = H_p(x^\varepsilon, c + v_x^\varepsilon(x^\varepsilon))$. Then $x(s) \in \mathcal{M}_\alpha^\varepsilon(c)$ and $v_x^\varepsilon(x(s)) = L_\xi(x(s), x'(s)) - c$ for all $s \in \mathbb{R}$.*

Proof. Let x^ε be an arbitrary point of $\mathcal{M}_\alpha^\varepsilon(c)$ for which there exists a minimizing curve $\gamma^\varepsilon : (-\infty, 0] \rightarrow \mathbb{T}^n$ and $\tau_j \rightarrow -\infty$ ($j \rightarrow \infty$) such that $\gamma^\varepsilon(\tau_j) \rightarrow x^\varepsilon$ ($j \rightarrow \infty$). Since

$$\gamma^{\varepsilon'}(\tau_j) = H_p(\gamma^\varepsilon(\tau_j), c + v_x^\varepsilon(\gamma^\varepsilon(\tau_j))) \Leftrightarrow v_x^\varepsilon(\gamma^\varepsilon(\tau_j)) = L_\xi(\gamma^\varepsilon(\tau_j), \gamma^{\varepsilon'}(\tau_j)) - c$$

is a bounded sequence, there exists a subsequence, still denoted by $\{\gamma^{\varepsilon'}(\tau_j)\}_j$, which converges to some $\xi^\varepsilon \in \mathbb{R}^n$. Let $x(s) : \mathbb{R} \rightarrow \mathbb{T}^n$ be the solution of (3.2) with the initial condition $x(0) = x^\varepsilon$, $x'(0) = \xi^\varepsilon$. It follows from the continuous dependence on initial conditions that we have for any $\alpha > 0$,

$$(\gamma^\varepsilon(\cdot + \tau_j), \gamma^{\varepsilon'}(\cdot + \tau_j)) \rightarrow (x(\cdot), x'(\cdot)) \text{ uniformly on } [-\alpha, \alpha] \text{ } (j \rightarrow \infty).$$

Hence, we see that $x(s) \in \mathcal{M}_\alpha^\varepsilon(c)$ for all $s \in \mathbb{R}$ by the definition of $\mathcal{M}_\alpha^\varepsilon(c)$. Since $\gamma^\varepsilon(\cdot + \alpha + \tau_j) : [-2\alpha, 0] \rightarrow \mathbb{T}^n$ is the unique minimizing curve for $v(\gamma^\varepsilon(\alpha + \tau_j))$ for each fixed large j , i.e.,

$$v^\varepsilon(\gamma^\varepsilon(\alpha + \tau_j)) = \int_{-2\alpha}^0 e^{\varepsilon s} (L(\gamma^\varepsilon(s + \alpha + \tau_j), \gamma^{\varepsilon'}(s + \alpha + \tau_j)) - c \cdot \gamma^{\varepsilon'}(s + \alpha + \tau_j) \\ + h(c)) ds + e^{-\varepsilon \cdot 2\alpha} v^\varepsilon(\gamma^\varepsilon(-\alpha + \tau_j)).$$

We obtain by letting $j \rightarrow \infty$,

$$\begin{aligned} v^\varepsilon(x(\alpha)) &= \int_{-2\alpha}^0 e^{\varepsilon s} (L(x(\alpha+s), x'(\alpha+s)) - c \cdot x'(\alpha+s) + h(c)) ds \\ &\quad + e^{-\varepsilon \cdot 2\alpha} v^\varepsilon(x(-\alpha)). \end{aligned}$$

Therefore, we see that $x(\alpha+s)$, $s \in [-2\alpha, 0]$ is a minimizing curve for $v^\varepsilon(x(\alpha))$, and hence that $v_x^\varepsilon(x(\alpha+s)) = L_\xi(x(\alpha+s), x'(\alpha+s)) - c$ for all $s \in [-2\alpha, 0)$. For $s = -\alpha$, we obtain $v_x^\varepsilon(x^\varepsilon) = L_\xi(x^\varepsilon, \xi^\varepsilon) - c$ and $\xi^\varepsilon = x'(0) = H_p(x^\varepsilon, c + v_x(x^\varepsilon))$. Continuity yields the same for each accumulating point $(x^\varepsilon, \xi^\varepsilon)$ of $\mathcal{M}_\alpha^\varepsilon(c)$. Noting that $\alpha > 0$ is arbitrary, we complete the proof. \square

The following sets are well-defined:

$$\begin{aligned} \tilde{\mathcal{M}}_\alpha^\varepsilon(c) &:= \{(x^\varepsilon, H_p(x^\varepsilon, c + v_x^\varepsilon(x^\varepsilon))) \mid x^\varepsilon \in \mathcal{M}_\alpha^\varepsilon(c)\}, \\ \tilde{\mathcal{M}}_\alpha^{\varepsilon*}(c) &:= \{(x^\varepsilon, c + v_x^\varepsilon(x^\varepsilon)) \mid x^\varepsilon \in \mathcal{M}_\alpha^\varepsilon(c)\}. \end{aligned}$$

Remark 3.5. The set $\mathcal{M}_\alpha^\varepsilon(c)$ is the closure of the family of all the curves $x(s)$ in Theorem 3.4, and

$$\tilde{\mathcal{M}}_\alpha^\varepsilon(c) = \overline{\{(x^\varepsilon, \xi^\varepsilon) \mid (x^\varepsilon, \xi^\varepsilon): \alpha\text{-limit point of } (\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)), \\ \gamma^\varepsilon: \text{minimizing curve of } v^\varepsilon\}}.$$

Here are properties of $\tilde{\mathcal{M}}_\alpha^\varepsilon(c)$ and $\tilde{\mathcal{M}}_\alpha^{\varepsilon*}(c)$:

- Theorem 3.6.** (i) For each point $x \in \mathbb{T}^n$, there exists $\xi \in \mathbb{R}^n$ such that $\phi_{L,c,\varepsilon}^s(x, \xi)$ falls into $\tilde{\mathcal{M}}_\alpha^\varepsilon(c)$ as $s \rightarrow -\infty$, i.e., any accumulating point of $\{\phi_{L,c,\varepsilon}^s(x, \xi)\}_{s \leq 0}$ belongs to $\tilde{\mathcal{M}}_\alpha^\varepsilon(c)$. If $v_x^\varepsilon(x)$ exists, $\xi = H_p(x, c + v_x^\varepsilon(x))$.
- (ii) $\tilde{\mathcal{M}}_\alpha^\varepsilon(c)$ is a $\phi_{L,c,\varepsilon}^s$ -invariant subset of $\mathbb{T}^n \times \mathbb{R}^n$, i.e., $\phi_{L,c,\varepsilon}^s(\tilde{\mathcal{M}}_\alpha^\varepsilon(c)) = \tilde{\mathcal{M}}_\alpha^\varepsilon(c)$ for all $s \in \mathbb{R}$.
- (iii) For each point $x \in \mathbb{T}^n$, there exists $p \in \mathbb{R}^n$ such that $\phi_{H,c,\varepsilon}^s(x, p)$ falls into $\tilde{\mathcal{M}}_\alpha^{\varepsilon*}(c)$ as $s \rightarrow -\infty$ along $\text{graph}(c + v_x^\varepsilon)$. If $v_x^\varepsilon(x)$ exists, $p = c + v_x^\varepsilon(x)$.
- (iv) $\tilde{\mathcal{M}}_\alpha^{\varepsilon*}(c)$ is a $\phi_{H,c,\varepsilon}^s$ -invariant subset of $\text{graph}(c + v_x^\varepsilon)$.

Proof is the same as that of Theorem 2.4.

3.3 Two types of minimizing measure

We observe measure-theoretical characterization of the discount problem and show relation between $\mathcal{M}_\alpha^\varepsilon(c)$, $\tilde{\mathcal{M}}_\alpha^\varepsilon(c)$ and minimizing measures. There are two types of minimizing measures obtained from (M1) $^\varepsilon$ and (M2) $^\varepsilon$ in Introduction. As is observed below, (M1) $^\varepsilon$ is a direct generalization of (M), and is well-related to $\mathcal{M}_\alpha^\varepsilon(c)$, $\tilde{\mathcal{M}}_\alpha^\varepsilon(c)$. On the other hand, (M2) $^\varepsilon$ is obtained naturally from (3.1) [7], or the nonlinear adjoint method as a solution of the adjoint PDE [22]. We will see that the minimizing measures of (M2) $^\varepsilon$ are different from those of (M1) $^\varepsilon$, namely the minimizing measures of (M2) $^\varepsilon$ are not

$\phi_{L,c,\varepsilon}^s$ -invariant in general and their support is not related to $\mathcal{M}_\alpha^\varepsilon(c)$, $\tilde{\mathcal{M}}_\alpha^\varepsilon(c)$, whereas both tend to some of the minimizing measures of (M) as $\varepsilon \rightarrow 0+$. In the selection problem [7], [22], the minimizing measures of (M2) $^\varepsilon$ are exploited.

Now we observe more on (M1) $^\varepsilon$. Let μ be arbitrary $\phi_{L,c,\varepsilon}^s$ -invariant probability measure. Then we have for any $(x, \xi) \in \mathbb{T}^n \times \mathbb{R}^n$,

$$(3.6) \quad v^\varepsilon(\pi \circ \phi_{L,c,\varepsilon}^0(x, \xi)) \leq \int_{-T}^0 e^{\varepsilon s} \{L^c(\phi_{L,c,\varepsilon}^s(x, \xi)) + h(c)\} ds + \varepsilon^{-\varepsilon T} v^\varepsilon(\pi \circ \phi_{L,c,\varepsilon}^{-T}(x, \xi)).$$

Integrating the inequality with μ over $\mathbb{T}^n \times \mathbb{R}^n$, we obtain

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} v^\varepsilon(x) d\mu \leq \frac{1 - e^{-\varepsilon T}}{\varepsilon} \left\{ \int_{\mathbb{T}^n \times \mathbb{R}^n} L^c(x, \xi) d\mu + h(c) \right\} + e^{-\varepsilon T} \int_{\mathbb{T}^n \times \mathbb{R}^n} v^\varepsilon(x) d\mu.$$

Letting $T \rightarrow \infty$, we have

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, \xi) - c \cdot \xi - \varepsilon v^\varepsilon(x) d\mu \geq -h(c).$$

Let $\gamma^\varepsilon(-\infty, 0] \rightarrow \mathbb{T}^n$ be a minimizing curve of v^ε . Note that $\{(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s))\}_{s \leq 0}$ is contained in a compact set $K \subset \mathbb{T}^n \times \mathbb{R}^n$ independent from ε and choice of γ^ε , because $\{v^\varepsilon\}_{\varepsilon > 0}$ is equi-Lipschitz continuous. Define the linear functional

$$\Psi_T^\varepsilon(\varphi) := \frac{1}{T} \int_{-T}^0 \varphi(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) ds, \quad \varphi \in C_c(\mathbb{T}^n \times \mathbb{R}^n, \mathbb{R}).$$

Then, by the Riesz representation theorem, there exists a probability measure μ_T on $\mathbb{T}^n \times \mathbb{R}^n$ such that

$$(3.7) \quad \Psi_T^\varepsilon(\varphi) = \int_{\mathbb{T}^n \times \mathbb{R}^n} \varphi(x, \xi) d\mu_T \quad \text{for all } \varphi \in C_c(\mathbb{T}^n \times \mathbb{R}^n, \mathbb{R}).$$

Since $\text{supp}(\mu_T)$ is contained in the compact set $K \subset \mathbb{T}^n \times \mathbb{R}^n$ for all $T > 0$ and $\varepsilon > 0$, we have $T_i \rightarrow \infty$ for which μ_{T_i} converges weakly to a probability measure μ^ε . We see that μ^ε is $\phi_{L,c,\varepsilon}^s$ -invariant. Furthermore, (3.7) with $\varphi(x, \xi) = \varepsilon v^\varepsilon(x)$ (in K , otherwise re-defined to be 0 continuously) and (1.2), as well as equivalence between (3.2) and (3.3) through the Legendre transform, yields

$$\begin{aligned} \int_{\mathbb{T}^n \times \mathbb{R}^n} \varepsilon v^\varepsilon(x) d\mu_{T_j} &= \frac{1}{T_j} \int_{-T_j}^0 \varepsilon v^\varepsilon(\gamma^\varepsilon(s)) ds \\ &= \frac{1}{T_j} \int_{-T_j}^0 -H(\gamma^\varepsilon(s), c + v_x^\varepsilon(\gamma^\varepsilon(s))) + h(c) ds \\ &= -\frac{1}{T_j} \int_{-T_j}^0 \gamma^{\varepsilon'}(s) \cdot \{c + v_x^\varepsilon(\gamma^\varepsilon(s))\} - L(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) ds + h(c) \\ &= -\frac{1}{T_j} \int_{-T_j}^0 \frac{d}{ds} \{v^\varepsilon(\gamma^\varepsilon(s))\} + \frac{1}{T_j} \int_{-T_j}^0 L^c(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) ds + h(c) \\ &= -\frac{v^\varepsilon(\gamma^\varepsilon(0)) - v^\varepsilon(\gamma^\varepsilon(-T_j))}{T_j} + \int_{\mathbb{T}^n \times \mathbb{R}^n} L^c(x, \xi) d\mu_{T_j} + h(c). \end{aligned}$$

Hence, we obtain with $j \rightarrow \infty$,

$$(3.8) \quad \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, \xi) - c \cdot \xi - \varepsilon v^\varepsilon(x) d\mu^\varepsilon = -h(c),$$

namely, μ^ε is a minimizing measure of $(M1)^\varepsilon$. It is clear that μ^ε satisfies the holonomic condition

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \psi_x(x) \cdot \xi d\mu^\varepsilon = 0 \quad \text{for all } \psi \in C^1(\mathbb{T}^n).$$

We examine the support of minimizing measures of $(M1)^\varepsilon$. Let μ^ε be an arbitrary minimizing measure of $(M1)^\varepsilon$. Since (3.6) holds for any $(x, \xi) \in \mathbb{T}^n \times \mathbb{R}^n$, the minimizing property (3.8) of μ^ε implies that we have

$$v^\varepsilon(\pi \circ \phi_{L,c,\varepsilon}^0(x, \xi)) = \int_{-T}^0 e^{\varepsilon s} \{L^c(\phi_{L,c,\varepsilon}^0(x, \xi)) + h(c)\} ds + \varepsilon^{-\varepsilon T} v^\varepsilon(\pi \circ \phi_{L,c,\varepsilon}^{-T}(x, \xi))$$

for each $(x, \xi) \in \text{supp}(\mu^\varepsilon)$. Hence, we have $\xi = H_p(x, c + v_x^\varepsilon(x))$ for each $(x, \xi) \in \text{supp}(\mu^\varepsilon)$, which means that $\text{supp}(\mu^\varepsilon)$ is contained in a compact set $K \subset \mathbb{T}^n \times \mathbb{R}^n$ for all $\varepsilon > 0$. Since $\text{supp}(\mu^\varepsilon) \subset K$ for all $\varepsilon > 0$, there exists a weakly convergent subsequence μ^{ε_j} , $\varepsilon_j \rightarrow 0+$ ($j \rightarrow \infty$), whose limit is a minimizing measure of (M) . Define the set for each $\varepsilon > 0$,

$$\tilde{\mathcal{M}}^\varepsilon(c) := \overline{\bigcup_{\mu^\varepsilon} \text{supp}(\mu^\varepsilon)}, \quad \mathcal{M}^\varepsilon(c) := \pi \tilde{\mathcal{M}}^\varepsilon(c),$$

where the union is taken with respect to all minimizing measures of $(M1)^\varepsilon$. Reasoning similar to the proof of Theorem 2.5 in Section 2 shows that each point $(x, \xi) \in \text{supp}(\mu^\varepsilon)$ itself is an α -limit point of $\phi_{L,c,\varepsilon}^s(x, \xi)$, where μ^ε is a minimizing measure of $(M1)^\varepsilon$. Therefore, we have the following inclusion:

Theorem 3.7. $\mathcal{M}^\varepsilon(c) \subset \mathcal{M}_\alpha^\varepsilon(c) \subset \mathbb{T}^n$, $\tilde{\mathcal{M}}^\varepsilon(c) \subset \tilde{\mathcal{M}}_\alpha^\varepsilon(c) \subset \mathbb{T}^n \times \mathbb{R}^n$.

Now we observe more on $(M2)^\varepsilon$. Let $x_0 \in \mathbb{T}^n$ be fixed and $\gamma^\varepsilon : (-\infty, 0] \rightarrow \mathbb{T}^n$ be a minimizing curve for $v^\varepsilon(x_0)$. Define the linear functional

$$\tilde{\Psi}_T^\varepsilon(\varphi) := \frac{\varepsilon}{1 - e^{-\varepsilon T}} \int_{-T}^0 e^{\varepsilon s} \varphi(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) ds, \quad \varphi \in C_c(\mathbb{T}^n \times \mathbb{R}^n, \mathbb{R}).$$

Then, by the Riesz representation theorem, there exists a probability measure $\tilde{\mu}_T$ on $\mathbb{T}^n \times \mathbb{R}^n$ such that

$$(3.9) \quad \tilde{\Psi}_T^\varepsilon(\varphi) = \int_{\mathbb{T}^n \times \mathbb{R}^n} \varphi(x, \xi) d\tilde{\mu}_T \quad \text{for all } \varphi \in C_c(\mathbb{T}^n \times \mathbb{R}^n, \mathbb{R}).$$

Since $\text{supp}(\tilde{\mu}_T)$ is contained in a compact set $K \subset \mathbb{T}^n \times \mathbb{R}^n$ for all $T > 0$ and $\varepsilon > 0$, we have $T_i \rightarrow \infty$ for which $\tilde{\mu}_{T_i}$ converges weakly to a probability measure $\tilde{\mu}^\varepsilon$ as $i \rightarrow \infty$. For $\varphi(x, \xi) = L^c(x, \xi) + h(c)$ (in K , otherwise re-defined to be 0 continuously), we have

$$\begin{aligned} \int_{\mathbb{T}^n \times \mathbb{R}^n} L^c(x, \xi) + h(c) d\tilde{\mu}_{T_i} &= \frac{\varepsilon}{1 - e^{-\varepsilon T}} \int_{-T_i}^0 e^{\varepsilon s} \{L^c(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) + h(c)\} ds \\ &= \frac{\varepsilon}{1 - e^{-\varepsilon T_i}} \{v^\varepsilon(x_0) - e^{-\varepsilon T_i} v^\varepsilon(\gamma^\varepsilon(-T_i))\}, \end{aligned}$$

and hence with $i \rightarrow \infty$,

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, \xi) - c \cdot \xi d\tilde{\mu}^\varepsilon = -h(c) + \varepsilon v^\varepsilon(x_0).$$

For $\varphi(x, \xi) = \psi_x(x) \cdot \xi$ (in K , otherwise re-defined to be 0 continuously) with $\psi \in C^1(\mathbb{T}^n)$, we have

$$\begin{aligned} \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi_x(x) \cdot \xi d\tilde{\mu}_{T_i} &= \frac{\varepsilon}{1 - e^{-\varepsilon T}} \int_{-T_i}^0 e^{\varepsilon s} \psi_x(\gamma^\varepsilon(s)) \cdot \gamma^{\varepsilon'}(s) ds \\ &= \frac{\varepsilon}{1 - e^{-\varepsilon T_i}} \int_{-T_i}^0 e^{\varepsilon s} \frac{d}{ds} \{\psi(\gamma^\varepsilon(s))\} \\ &= \frac{\varepsilon}{1 - e^{-\varepsilon T_i}} \left\{ e^{\varepsilon s} \psi(\gamma^\varepsilon(s)) \Big|_{-T_i}^0 - \varepsilon \int_{-T_i}^0 e^{\varepsilon s} \psi(\gamma^\varepsilon(s)) ds \right\} \\ &= \frac{\varepsilon}{1 - e^{-\varepsilon T_i}} \left\{ \psi(x_0) - e^{-\varepsilon T_i} \psi(\gamma^\varepsilon(-T_i)) \right\} - \varepsilon \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x) d\tilde{\mu}_{T_i}. \end{aligned}$$

Letting $i \rightarrow \infty$, we see that $\tilde{\mu}^\varepsilon$ satisfies (1.4). In order to see that $\tilde{\mu}^\varepsilon$ is a minimizing measure of $(M2)^\varepsilon$, we check

$$(3.10) \quad \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, \xi) - c \cdot \xi d\mu \geq -h(c) + \varepsilon v^\varepsilon(x_0)$$

to be true for any probability measure μ satisfying (1.4). Let u^η denote $v^\varepsilon * \rho^\eta$, where ρ^η is the standard mollifier on \mathbb{T}^n with the parameter $\eta \rightarrow 0+$. Note that $u_x^\eta = v_x^\varepsilon * \rho^\eta$ and $u^\eta \rightarrow v^\varepsilon$ uniformly as $\eta \rightarrow 0+$. Then, the Legendre transform and Jensen's inequality imply that

$$\begin{aligned} (3.11) \quad \varepsilon u^\eta(x) + \xi \cdot (c + u_x^\eta(x)) - L(x, \xi) &\leq \varepsilon u^\eta(x) + H(x, c + u_x^\eta(x)) \\ &\leq \{\varepsilon v^\varepsilon(\cdot) + H(x, c + v_x^\varepsilon(\cdot))\} * \rho^\eta(x) \\ &= \{\varepsilon v^\varepsilon(\cdot) + H(\cdot, c + v_x^\varepsilon(\cdot))\} * \rho^\eta(x) + O(\eta) \\ &= h(c) + O(\eta) \quad \text{for any } (x, \xi) \in \mathbb{T}^n \times \mathbb{R}^n. \end{aligned}$$

Hence, we have with (1.4),

$$\begin{aligned} h(c) + O(\eta) &\geq \varepsilon \int_{\mathbb{T}^n \times \mathbb{R}^n} u^\eta(x) d\mu + \int_{\mathbb{T}^n \times \mathbb{R}^n} \xi \cdot (c + u_x^\eta(x)) - L(x, \xi) d\mu \\ &= \varepsilon \int_{\mathbb{T}^n \times \mathbb{R}^n} u^\eta(x) d\mu + \int_{\mathbb{T}^n \times \mathbb{R}^n} \xi \cdot u_x^\eta(x) - \int_{\mathbb{T}^n \times \mathbb{R}^n} L^c(x, \xi) d\mu \\ &= \varepsilon u^\eta(x_0) - \int_{\mathbb{T}^n \times \mathbb{R}^n} L^c(x, \xi) d\mu. \end{aligned}$$

Letting $\eta \rightarrow 0+$, we have (3.10).

We examine the support of each minimizing measures of $(M2)^\varepsilon$. Let $\tilde{\mu}^\varepsilon$ be an arbitrary minimizing measure of $(M2)^\varepsilon$. It follows from (3.11) and the minimizing property of $\tilde{\mu}^\varepsilon$ that for each $(x, \xi) \in \text{supp}(\tilde{\mu}^\varepsilon)$ we have

$$\varepsilon u^\eta(x) + \xi \cdot (c + u_x^\eta(x)) - L(x, \xi) - h(c) \rightarrow 0 \text{ as } \eta \rightarrow 0+.$$

Since $u_x^\eta(\cdot)$ is uniformly bounded in \mathbb{T}^n , we have with each accumulating point p of $\{u_x^\eta(x)\}_{\eta>0}$, which is bounded,

$$\varepsilon v^\varepsilon(x) + \xi \cdot (c + p) - L(x, \xi) - h(c) = 0 \quad \text{for each } (x, \xi) \in \text{supp}(\tilde{\mu}^\varepsilon).$$

Hence, ξ is bounded independently of x and ε , namely there exists a compact set $K \subset \mathbb{T}^n \times \mathbb{R}^n$ such that $\text{supp}(\tilde{\mu}^\varepsilon) \subset K$ for any $\varepsilon > 0$. If $v_x^\varepsilon(x)$ exists, we have $\xi = H_p(x, c + v_x^\varepsilon(x))$.

Remark 3.8. *A minimizing measure $\tilde{\mu}^\varepsilon$ of $(M2)^\varepsilon$ is NOT $\phi_{L,c,\varepsilon}^s$ -invariant in general. Furthermore, the viscosity solution to (1.2) is not differentiable at every point of the support of $\tilde{\mu}^\varepsilon$ in general.*

In order to check Remark 3.8, suppose that v^ε is not differentiable at $x = x_0$. Then, a minimizing curve γ^ε for $v^\varepsilon(x_0)$ does not have x_0 as its α -limit point. We will see that the support of the minimizing measure $\tilde{\mu}^\varepsilon$ induced by γ^ε through (3.9) contains (x_0, ξ_0) with $\xi_0 := \gamma^{\varepsilon'}(0)$. Let B_δ be a δ -ball of (x_0, ξ_0) and φ be a positive function supported inside B_δ . There exists $\tau_\delta^- > 0$ such that $\phi_{L,c,\varepsilon}^s(x_0, \xi_0)$ first touches ∂B_δ at $s = \tau_\delta^-$ for $s < 0$. Since x_0 is not an α -limit point of γ^ε , $\phi_{L,c,\varepsilon}^s(x_0, \xi_0)$ never comes back to B_δ for $s < \tau_\delta^-$. Hence, we have

$$\begin{aligned} \int_{\mathbb{T}^n \times \mathbb{R}^n} \varphi(x, \xi) d\tilde{\mu}^\varepsilon &= \lim_{i \rightarrow \infty} \frac{\varepsilon}{1 - e^{-\varepsilon T_i}} \int_{-T_i}^0 e^{\varepsilon s} \varphi(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) ds \\ &= \varepsilon \int_{\tau_\delta^-}^0 e^{\varepsilon s} \varphi(\gamma^\varepsilon(s), \gamma^{\varepsilon'}(s)) ds > 0. \end{aligned}$$

Since $\delta > 0$ is arbitrary, we have $(x_0, \xi_0) \in \text{supp}(\tilde{\mu}^\varepsilon)$. We see that $\tilde{\mu}^\varepsilon$ is not $\phi_{L,c,\varepsilon}^s$ -invariant. Let $\tau_\delta^+ > 0$ be such that $\phi_{L,c,\varepsilon}^s(x_0, \xi_0)$ first touches ∂B_δ at $s = \tau_\delta^+$ for $s > 0$. Fix $\tau > 0$. Take $\delta > 0$ so small that $\phi^{-\tau}(B_\delta) \cap B_\delta = \emptyset$. Then $\phi_{L,c,\varepsilon}^s(x_0, \xi_0)$ stays in $\phi^{-\tau}(B_\delta)$ only within $s \in (-\tau - \tau_\delta^-, -\tau + \tau_\delta^+)$ for $\delta \rightarrow 0+$. Hence, with the indicator function χ_A for each $A \subset \mathbb{T}^n \times \mathbb{R}^n$ and a standard mollifier ρ^η on $\mathbb{T}^n \times \mathbb{R}^n$, we have

$$\begin{aligned} \tilde{\mu}^\varepsilon(B_\delta) &= \lim_{i \rightarrow \infty} \lim_{\eta \rightarrow 0+} \int_{\mathbb{T}^n \times \mathbb{R}^n} (\chi_{B_\delta} * \rho^\eta)(x, \xi) d\tilde{\mu}_{T_i} \\ &= \lim_{i \rightarrow \infty} \lim_{\eta \rightarrow 0+} \frac{\varepsilon}{1 - e^{-\varepsilon T_i}} \int_{-T_i}^0 e^{\varepsilon s} (\chi_{B_\delta} * \rho^\eta)(\gamma^*(s), \gamma^{*'}(s)) ds \\ &= 1 - e^{-\varepsilon \tau_\delta^-}, \\ \tilde{\mu}^\varepsilon(\phi_{L,c,\varepsilon}^{-\tau}(B_\delta)) &= \lim_{i \rightarrow \infty} \lim_{\eta \rightarrow 0+} \int_{\mathbb{T}^n \times \mathbb{R}^n} (\chi_{\phi_{L,c,\varepsilon}^{-\tau}(B_\delta)} * \rho^\eta)(x, \xi) d\tilde{\mu}_{T_i} \\ &= \lim_{i \rightarrow \infty} \lim_{\eta \rightarrow 0+} \frac{\varepsilon}{1 - e^{-\varepsilon T_i}} \int_{-T_i}^0 \left(e^{\varepsilon s} \chi_{\phi_{L,c,\varepsilon}^{-\tau}(B_\delta)} * \rho^\eta \right)(\gamma^*(s), \gamma^{*'}(s)) ds \\ &= e^{-\varepsilon(\tau - \tau_\delta^+)} - e^{-\varepsilon(\tau + \tau_\delta^-)}, \end{aligned}$$

which are not identical. This implies also that v^ε is not necessarily differentiable at every point of $\pi \text{supp}(\tilde{\mu}^\varepsilon)$, e.g., at x_0 in the above example.

Define the sets for each $\varepsilon > 0$,

$$\tilde{\mathcal{M}}^\varepsilon(c) := \overline{\bigcup_{\tilde{\mu}^\varepsilon} \text{supp}(\tilde{\mu}^\varepsilon)}, \quad \hat{\mathcal{M}}^\varepsilon(c) := \pi \tilde{\mathcal{M}}^\varepsilon(c),$$

where the union is taken with respect to all minimizing measures of $(M2)^\varepsilon$ for all $x_0 \in \mathbb{T}^n$. Unlike $(M1)^\varepsilon$, minimizers of $(M2)^\varepsilon$ do not provide information on α -limit points, and in particular,

$$\tilde{\mathcal{M}}^\varepsilon(c) \not\subset \tilde{\mathcal{M}}_\alpha^\varepsilon(c), \quad \hat{\mathcal{M}}^\varepsilon(c) \not\subset \mathcal{M}_\alpha^\varepsilon(c) \text{ in general.}$$

Nevertheless, since each minimizing measure of $(M2)^\varepsilon$ is supported on a compact set K for any $\varepsilon > 0$, there exists weakly convergent sequence $\tilde{\mu}^{\varepsilon_i}$ with $\varepsilon_i \rightarrow 0+$ ($i \rightarrow \infty$). Since (1.4) yields the holonomic condition (2.8) for $\varepsilon \rightarrow 0+$, the limit of $\tilde{\mu}^{\varepsilon_i}$ is a minimizing measure of (M) .

4 Error estimate in vanishing discount process

We show error estimates between viscosity solutions of (1.1) and (1.2) in the following two cases:

(C1) (1.1) with $H(x, p) = \frac{1}{2}p^2 - F(x) : \mathbb{T} \times \mathbb{R} \rightarrow \mathbb{R}$ whose corresponding Hamiltonian system possesses hyperbolic stationary solutions.

(C2) (1.1) with a C^2 -solution which generates a KAM n -torus.

In the first case, we could see the advantage to introducing the families of α -limit points $\mathcal{M}_\alpha(c)$ and $\mathcal{M}_\alpha^\varepsilon(c)$, where asymptotics of the discount dynamical systems including $\mathcal{M}_\alpha^\varepsilon(c)$ for $\varepsilon \rightarrow 0+$ is also made clear.

We refer to [7, Theorem 1.1], [22, Theorem 1.1] for the general convergence result in the vanishing discount process. Under (H1)-(H3), there exists a viscosity solution v^* to (1.1) such that

$$v^\varepsilon \rightarrow v^* \text{ uniformly on } \mathbb{T}^n \text{ as } \varepsilon \rightarrow 0+.$$

Moreover, v^* is characterized by

(4.1) $v^*(x) = \sup\{v(x) \mid v \in \text{Lip}(\mathbb{T}^n), \text{ viscosity subsolution to (1.1) such that}$

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} v \, d\mu \leq 0 \text{ for all Mather measures } \mu \text{ of (M)} \}.$$

We remark that a criterion in this characterization is first observed in [15, Corollary 4]. We also refer to [5] for an L^2 -estimate with respect to a discount Mather measure between the gradient of the viscosity solution to (1.2) and that of the mollified function of a viscosity solution to (1.1) on the support of the measure.

4.1 Error estimate in (C1)

We examine asymptotics of the discount problem for $\varepsilon \rightarrow 0+$ with

$$H(x, p) := \frac{1}{2}p^2 - F(x) : \mathbb{T} \times \mathbb{R} \rightarrow \mathbb{R}, \quad F \in C^2(\mathbb{T}), \quad F(x) \geq 0 \quad \text{in } \mathbb{T}.$$

Suppose that F has a finite number of zero points $0 \leq x_1 < x_2 < \dots < x_I < 1$ at which $F_{xx}(x_i) > 0$ for all $i = 1, 2, \dots, I$. Since $F_x(x_i) = 0$, the points $(x_i, 0)$ are hyperbolic

stationary solutions of the Hamiltonian system (2.4), whose stable/unstable manifolds form a separatrix

$$S_{\pm} := \{(x, \pm\sqrt{2F(x)}) \mid x \in \mathbb{T}\}.$$

Note that the slope of S_{\pm} at x_i is $\pm\sqrt{F_{xx}(x_i)}$. Set $c_{\pm} := \pm \int_{\mathbb{T}} \sqrt{2F(x)} dx$. Due to such a structure of the Hamiltonian system, we see that the Hamilton-Jacobi equation (1.1) has the following properties:

- $h(c) \equiv 0$ for $c \in [c_-, c_+]$.
- $\mathcal{M}_{\alpha}(v; c) = \{x_1, x_2, \dots, x_I\}$ for each $c \in [c_-, c_+]$ and viscosity solution v of (1.1).
- Let $c \in (c_-, c_+)$ and v be a viscosity solution of (1.1). Then the graph of $c + v_x$ consists of parts of S_{\pm} and has at least one point of discontinuity in such a way that the jump is only from S_+ to S_- (the entropy condition).
- Let $c \in (c_-, c_+)$. If a function $u(x)$ is such that $\int_{\mathbb{T}} c + u(x) dx = c$ and the graph of $c + u(x)$ consists of parts of S_{\pm} , and has discontinuity only from S_+ to S_- , then the primitive function of u is a viscosity solution of (1.1).
- If $I \geq 2$, such functions as u may have 2 to I points of singularity and may exist uncountably many (shift the position of discontinuity right or left keeping $\int_{\mathbb{T}} c + u(x) dx = c$). This means that there are uncountably many viscosity solutions of (1.1) beyond constant difference.

Note carefully that the points between x_{i-1} and x_i at which $F_x = 0$ also give stationary solutions of (2.4), where they are elliptic if $F_{xx} < 0$ or hyperbolic if $F_{xx} > 0$.

We study the limit process $\varepsilon \rightarrow 0+$ of (1.2) for $c \in (c_-, c_+)$, exploiting dynamical properties of (3.3) in the present situation. The first task is to specify $\mathcal{M}_{\alpha}^{\varepsilon}(c)$. Here, (3.3) is of the form

$$(4.2) \quad x'(s) = p(s), \quad p'(s) = F_x(x(s)) + c\varepsilon - \varepsilon p(s).$$

For each small $\varepsilon > 0$, (4.2) has hyperbolic stationary solutions $(x_i^{c,\varepsilon}, 0)$ which are $O(\varepsilon)$ -close to $(x_i, 0)$ for $i = 1, 2, \dots, I$; namely, $x_i^{c,\varepsilon}$ is the value of the implicit function derived from $F_x(x) + c\varepsilon = 0$ near $(\varepsilon, x) = (0, x_i)$. Therefore we have the local stable/unstable manifolds of $(x_i^{\varepsilon}, 0)$, which tend to the exact local stable/unstable manifolds of $(x_i, 0)$ as $\varepsilon \rightarrow 0$. Note carefully that the geometrical structure of the phase space of (4.2) is not equivalent to that of (1.1) no matter how small $\varepsilon > 0$ is, i.e., the stable/unstable manifolds of $(x_i^{\varepsilon}, 0)$ (the extension of the local ones by $\phi_{H,c,\varepsilon}^s$ for $s \in \mathbb{R}$) do not connect each other and hence do not form a separatrix like S_{\pm} in general, where some of the stationary solutions of (2.4) may change to be asymptotically stable due to non-Hamiltonian perturbation in (4.2). However the stable/unstable manifolds cannot be transversal in our autonomous 1-dimensional setting.

Theorem 4.1. *Set $\Gamma^{c,\varepsilon} := \{x_1^{c,\varepsilon}, \dots, x_I^{c,\varepsilon}\}$, where $x_i^{c,\varepsilon}$ is the value of the implicit function derived from $F_x(x) + c\varepsilon = 0$ near $(\varepsilon, x) = (0, x_i)$. For each sufficiently small $\varepsilon > 0$, it holds that*

- (i) $c + v_x^{\varepsilon}(x) = 0$ on $\mathcal{M}_{\alpha}^{\varepsilon}(c)$ for each $c \in (c_-, c_+)$.

- (ii) $\mathcal{M}_\alpha^\varepsilon(c) \subset \Gamma^{c,\varepsilon}$ for each $c \in (c_-, c_+)$.
- (iii) $|v^\varepsilon| \leq \beta\varepsilon$ on $\mathcal{M}_\alpha^\varepsilon(c)$ for each $c \in (c_-, c_+)$, where $\beta > 0$ is a constant independent of c and ε .
- (iv) There exists $c_+^* \in (0, c_+]$ and $c_-^* \in [c_-, 0)$ such that $\mathcal{M}_\alpha^\varepsilon(c) = \Gamma^{c,\varepsilon}$ for all $c \in [c_-^*, c_+^*]$.

Proof. (i) Fix $c \in (c_-, c_+)$. Suppose that there exists a sequence $\varepsilon_i \rightarrow 0+$ ($i \rightarrow \infty$) for which $\mathcal{M}_\alpha^{\varepsilon_i}(c)$ contains a point x^* such that $c + v_x^{\varepsilon_i}(x^*) \neq 0$ (v^{ε_i} is differentiable on $\mathcal{M}_\alpha^{\varepsilon_i}(c)$, as shown in Section 3). Let $\gamma^{\varepsilon_i}(s) : (-\infty, 0] \rightarrow \mathbb{T}$ be a minimizing curve that has x^* as its α -limit point. Then $(x(s), p(s)) := (\gamma^{\varepsilon_i}(s), c + v_x^{\varepsilon_i}(\gamma^{\varepsilon_i}(s)))$, which solves (4.2) for $s \leq 0$, is not on any unstable manifolds of (4.2) nor in the region that is absorbed by asymptotically unstable stationary solutions of (4.2), because otherwise we necessarily have $\lim_{s \rightarrow -\infty} (x(s), p(s)) = (x^*, 0) = (x^*, c + v_x^{\varepsilon_i}(x^*))$. Moreover $(x(s), p(s))$ cannot touch the line $p = 0$ for $s < 0$. If not, $(x(s), p(s))$ touches the line $p = 0$ away from any stationary solution of (4.2) at some $s = s_0$. Then $x''(s_0) = p'(s_0) = F_x(x(s_0)) + c\varepsilon_i \neq 0$, $x'(s_0) = 0$ and $x(s)$ takes a local maximum or minimum at $s = s_0$. Hence $p(s)$ changes its sign at $s = s_0$ and $\{(x(s), p(s)) \mid s \in [s_0 - \delta, s_0 + \delta]\}$ cannot be a single-valued graph defined on $\{x(s) \mid s \in [s_0 - \delta, s_0 + \delta]\}$. However $(x(s), p(s))$ must be on $\text{graph}(c + v_x^{\varepsilon_i})$ for $s < 0$, and we reach a contradiction. Hence, $\{(x(s), p(s)) \mid s < 0\} \subset \text{graph}(c + v_x^{\varepsilon_i})$ must be strictly above or below the line $p = 0$, where $x(s)$ moves over \mathbb{T} for $s < 0$.

Therefore, we conclude that $\text{graph}(c + v_x^{\varepsilon_i})$ is smooth and strictly above or below the line $p = 0$ for all i . Then, we have a subsequence of $c + v_x^{\varepsilon_i}$ that converges to $c + v_x$ pointwise a.e. as $i \rightarrow \infty$ (we do not yet assert that the whole sequence is convergent), where v is a viscosity solution. $\text{graph}(c + v_x)$ is above or below the line $p = 0$, which is only possible for $c = c_\pm$. Thus we obtain the conclusion.

(ii) Fix $c \in (c_-, c_+)$. Let $\varepsilon_0 > 0$ and $\delta > 0$ be such that the implicit function $x = x_i^{c,\varepsilon} : [-\varepsilon_0, \varepsilon_0] \rightarrow [x_i - \delta, x_i + \delta]$ is one to one and onto for all $i = 1, \dots, I$. Since $F(x) = 0$ only for $x = x_1, \dots, x_I$, there exists $\eta > 0$ such that $F(x) > \eta$ for all $x \notin [x_i - \delta, x_i + \delta]$, $i = 1, \dots, I$. Let x^* be an arbitrary point of $\mathcal{M}_\alpha^\varepsilon(c)$. Since $c + v_x^\varepsilon(x^*) = 0$ due to (i) and v^ε is uniformly bounded, it follows from (1.2) that for sufficiently small $\varepsilon > 0$,

$$F(x^*) = \varepsilon v^\varepsilon(x^*) \leq \eta.$$

Hence, $x^* \in [x_i - \delta, x_i + \delta]$ for some i . Since $(x^*, 0)$ is a stationary solution of (4.2), we have $F_x(x^*) + c\varepsilon = 0$, which implies $x^* = x_i^{c,\varepsilon}$.

(iii) In what follows, $\beta_j > 0$ are some constants independent of ε and c . Let x^* be an arbitrary point of $\mathcal{M}_\alpha^\varepsilon(c)$. Since x^* is equal to $x_i^{c,\varepsilon}$ for some i due to (ii), we have $|x^* - x_i| \leq \beta_1\varepsilon$. Hence, we observe with (i) and (1.2) that

$$\varepsilon v^\varepsilon(x^*) = F(x^*) = F(x^*) - F(x_i) = F_x(x_i) + \frac{1}{2}F_{xx}(x_i + \theta(x^* - x_i))(x^* - x_i)^2,$$

where $\theta \in (0, 1)$. Since $F_x(x_i) = 0$, we conclude that $|v^\varepsilon(x^*)| \leq \beta_2\varepsilon$.

(iv) Suppose that $\Gamma^{c,\varepsilon} \setminus \mathcal{M}_\alpha^\varepsilon(c)$ is non-empty. Let x^* be an arbitrary point of $\Gamma^{c,\varepsilon} \setminus \mathcal{M}_\alpha^\varepsilon(c)$. Since x^* is equal to $x_i^{c,\varepsilon}$ for some i , we have $|x^* - x_i| \leq \beta_1\varepsilon$. By (1.2), we observe

that if v^ε is differentiable at x^*

$$\begin{aligned}\varepsilon v^\varepsilon(x^*) &= F(x^*) - \frac{1}{2}(c + v_x^\varepsilon(x^*))^2 \\ &\leq F(x^*) = F_x(x_i) + \frac{1}{2}F_{xx}(x_i + \theta(x^* - x_i))(x^* - x_i)^2 \leq \beta_2\varepsilon^2\end{aligned}$$

for some $\theta \in (0, 1)$, and hence we have

$$(4.3) \quad v^\varepsilon(x^*) \leq \beta_2\varepsilon.$$

If v^ε is not differentiable at x^* , take a point of differentiability of v^ε near x^* and use continuity of v^ε to obtain (4.3).

Since $x^* \notin \mathcal{M}_\alpha^\varepsilon(c)$, the minimizing curve $\gamma^\varepsilon : (-\infty, 0] \rightarrow \mathbb{R}$ (we do not take “mod 1” now) starting from a point x of differentiability of v^ε near x^* is such that $(\gamma^\varepsilon(s), c + v_x^\varepsilon(\gamma^\varepsilon(s)))$ stays on the unstable manifold of a certain stationary solution $(x_j^{c,\varepsilon}, 0)$ of (4.2) for $s \leq 0$, where $x_j^{c,\varepsilon} \in \mathcal{M}_\alpha^\varepsilon(c)$. Due to the argument in the proof of (i), $(\gamma^\varepsilon(s), c + v_x^\varepsilon(\gamma^\varepsilon(s)))$ cannot touch the line $p = 0$ for $s \leq 0$. If $c + v_x^\varepsilon(x) > 0$ (resp., $c + v_x^\varepsilon(x) < 0$), then $(\gamma^\varepsilon(s), c + v_x^\varepsilon(\gamma^\varepsilon(s)))$ is above (resp., below) the line $p = 0$ and falls into $(x_j^{c,\varepsilon} + m, 0)$ with $m = 0$ or -1 from the right (resp., $m = 0$ or 1 from the left) as $s \rightarrow -\infty$. Therefore, sending x to x^* , we see that

$$\begin{aligned}\beta_3 &< \int_{x_j^{c,\varepsilon}+m}^{x^*} c + v_x^\varepsilon(x) dx = c(x^* - x_j^{c,\varepsilon} - m) + v(x^*) - v^\varepsilon(x_j^{c,\varepsilon}) \\ (\text{resp., } -\beta_3 &> \int_{x^*}^{x_j^{c,\varepsilon}+m} c + v_x^\varepsilon(x) dx = c(x_j^{c,\varepsilon} + m - x^*) + v^\varepsilon(x_j^{c,\varepsilon}) - v^\varepsilon(x^*)),\end{aligned}$$

where $\beta_3 > 0$ is a number smaller than the area trimmed by the line $x = x^*$ and the unstable manifold of $(x_j^{0,\varepsilon} + m, 0)$ within $x \in [x_j^{0,\varepsilon} + m, x^*]$ (resp., within $x \in [x^*, x_j^{0,\varepsilon} + m]$). Here β_3 is of $O(1)$ for $\varepsilon \rightarrow 0$. Since $|v^\varepsilon(x_j^{c,\varepsilon})| \leq \beta\varepsilon$ due to (iii), we have

$$\begin{aligned}\beta_2\varepsilon &\geq v^\varepsilon(x^*) > \beta_3 - \beta\varepsilon - c(x^* - x_j^{c,\varepsilon} - m) \\ (\text{resp., } \beta_2\varepsilon &\geq v^\varepsilon(x^*) > \beta_3 - \beta\varepsilon + c(x_j^{c,\varepsilon} + m - x^*)).\end{aligned}$$

If $|c|$ is less than a certain value, we reach a contradiction. \square

Remark 4.2. $\mathcal{M}_\alpha^\varepsilon(c) = \Gamma^{c,\varepsilon}$ does not hold for every $c \in (c_-, c_+)$ in general, which implies that not every Mather measure is obtained through $(M1)^\varepsilon$ in the vanishing discount process.

For more detailed discussion in regards to Remark 4.2, let us consider the following example: $F \in C^2(\mathbb{T})$, $F(x) \geq 0$, $F(x) = 0$ for $x \in \{0, 1/2\}$, $c_\pm := \pm \int_{\mathbb{T}} \sqrt{2F(x)} dx = \pm 2$, $S_1 := \int_0^{1/2} \sqrt{2F(x)} dx = 1/2$. Set $x_1 = 0$ and $x_2 = 1/2$. Then, for each sufficiently small $\varepsilon > 0$, we have

$$\mathcal{M}_\alpha^\varepsilon(c) = \{x_1^{c,\varepsilon}\} \text{ for } c \in (1, 2), \quad \mathcal{M}_\alpha^\varepsilon(c) = \{x_2^{c,\varepsilon}\} \text{ for } c \in (-2, -1).$$

In fact, let $c \in (1, 2)$ (resp., $c \in (-2, -1)$) and suppose that there exists $\varepsilon_j \rightarrow 0$ such that $x_2^{c,\varepsilon_j} \in \mathcal{M}_\alpha^{\varepsilon_j}(c)$ (resp., $x_1^{c,\varepsilon_j} \in \mathcal{M}_\alpha^{\varepsilon_j}(c)$) for all j . By (3.1) with $\gamma(s) \equiv x_1, x_2$ and letting $T \rightarrow \infty$, we have $v^{\varepsilon_j}(x_1) \leq 0$, $v^{\varepsilon_j}(x_2) \leq 0$. Hence, an accumulating point v^* of

$\{v^{\varepsilon_j}\}$ satisfies $v^*(x_1) \leq 0$, $v^*(x_2) \leq 0$. In particular, by (iii) of Theorem 4.1, we have $O(\varepsilon_j) = v^{\varepsilon_j}(x_2^{c, \varepsilon_j}) = v^{\varepsilon_j}(x_2) + O(\varepsilon_j)$ (resp., $O(\varepsilon_j) = v^{\varepsilon_j}(x_1^{c, \varepsilon_j}) = v^{\varepsilon_j}(x_1) + O(\varepsilon_j)$), and hence $v^*(x_2) = 0$ (resp., $v^*(x_1) = 0$). Since $\text{graph}(c + v_x^*)$ consists of parts of the graph of $\pm\sqrt{2F(x)}$ with jump down discontinuity, we see that

$$\begin{aligned} \frac{1}{2} = S_1 &\geq \left| \int_{x_1}^{x_2} c + v_x^*(x) dx \right| = |c(x_2 - x_1) + v^*(x_2) - v^*(x_1)| = \left| \frac{c}{2} - v^*(x_1) \right| \geq \frac{c}{2}, \\ (\text{resp.}, \frac{1}{2} &\geq \left| \int_{x_1}^{x_2} c + v_x^*(x) dx \right| = |c(x_2 - x_1) + v^*(x_2) - v^*(x_1)| = \left| \frac{c}{2} + v^*(x_2) \right| \geq -\frac{c}{2}). \end{aligned}$$

If $c > 1$ (resp., $c < -1$), we reach a contradiction.

In this example, Mather measures are Dirac measures supported by x_1 and x_2 . It is worth emphasizing that we can clearly see the selection of Mather measures by $(M1)^\varepsilon$ depending on c .

Notice here that the selection of viscosity solutions and Mather measures in the vanishing discount process is totally different from that in the vanishing viscosity process. Indeed, if we assume $0 < F_{xx}(x_1) < F_{xx}(x_2)$, then for each $c \in (-2, 2)$ the vanishing viscosity method selects the viscosity solution \tilde{v} up to constant such that $\text{graph}(c + \tilde{v}_x)$ contains the local unstable manifold of $(x_i, 0)$ in a neighborhood of x_i only for $i = 1$, which has only one point of discontinuity [18], [3], [2] (cf., if $c = 0$, for instance, $\text{graph}(c + v_x^*)$ has two points of discontinuity due to $v^*(x_1) = v^*(x_2) = 0$, which can also be observed by the formula (4.1)). We can also see a difference of the selection of Mather measures associated with the vanishing viscosity process (see [1] for details).

Now it is easy to observe that

- For $c \in [c_-^*, c_+^*]$, $v^\varepsilon(x) = O(\varepsilon)$ on $\{x_1, \dots, x_I\}$ for any $\varepsilon > 0$ and any accumulating point v^* of $\{v^\varepsilon\}_{\varepsilon>0}$ vanishes on $\mathcal{M}_\alpha(c) = \{x_1, \dots, x_I\}$. Such v^* is unique because of (iv) of Theorem 2.2 and therefore $v^\varepsilon \rightarrow v^*$ as $\varepsilon \rightarrow 0$ in the whole sequence. This is a specific example of the general convergence result in [7] and [22].
- There may exist $c \in (c_-, c_+)$ for which an accumulating point v^* of $\{v^\varepsilon\}_{\varepsilon>0}$ does not vanishes on $\{x_1, \dots, x_I\}$.

Next, we obtain error estimates between v^ε and $v^* = \lim_{\varepsilon \rightarrow 0+} v^\varepsilon$. Our strategy is to estimate the time for each minimizing curve to fall into $\mathcal{M}_\alpha^\varepsilon(c)$, $\mathcal{M}_\alpha(c)$ along the unstable manifolds and to compare the value functions (2.1) and (3.1). In what follows, $x_{I+k} := x_k + 1$ and each interval $K \subset \mathbb{R}$ is identified with $K \bmod 1 := \{x \bmod 1 \mid x \in K\}$.

Let \bar{x}_i be the midpoint of the interval $[x_i, x_{i+1}]$, $i = 1, 2, \dots, I$. Since the slope of the local stable/unstable manifold of each $(x_i, 0)$, given by the graph of $\pm\sqrt{2F(x)}$, is strictly away from 0 at $(x_i, 0)$, there exists $b > 0$ such that

$$\sqrt{2F(x)} \geq b(x - x_i) \text{ on } [x_i, \bar{x}_i], \quad \sqrt{2F(x)} \geq -b(x - x_{i+1}) \text{ on } [\bar{x}_i, x_{i+1}].$$

Hence, each trajectory $(x(s), p(s))$ on the unstable manifold of $(x_i, 0)$ with $x(0) \in (x_i, x_{i+1})$ satisfies the estimate

$$\text{if } x(0) \in (\bar{x}_i, x_{i+1}],$$

$$x'(t) = p(t) \geq -b(x(t) - x_{i+1}) \text{ and } x_{i+1} - x(-\tau) \geq (x_{i+1} - x(0))e^{b\tau},$$

$$\text{if } x(0) \in [x_i, \bar{x}_i],$$

$$x'(t) = p(t) \geq b(x(t) - x_i) \text{ and } 0 < x(-\tau) - x_i \leq (x(0) - x_i)e^{-b\tau}.$$

Therefore, for each $\delta > 0$, the time $\tau > 0$ for which the backward trajectory with $x(0) \in (\bar{x}_i, x_{i+1} - \delta]$ reaches the midpoint $(\bar{x}_i, \sqrt{2F(\bar{x}_i)})$ is estimated as

$$(4.4) \quad \tau \leq \frac{1}{b} \log \frac{1}{\delta}.$$

The same estimate holds for the time $\tau > 0$ for which the backward trajectory with $x(0) \in (\bar{x}_i + \delta, \bar{x}_i]$ reaches the δ -cylinder $[x_i - \delta, x_i + \delta] \times \mathbb{R}$. In this way, we see that each trajectory on the stable/unstable manifolds of $(x_i, 0)$ starting away from δ -cylinders $[x_i - \delta, x_i + \delta] \times \mathbb{R}$ ($i = 1, \dots, I$) reaches a δ -cylinder $[x_j - \delta, x_j + \delta] \times \mathbb{R}$ ($j = i - 1, i$ or $i + 1$) taking at most time $T > 0$:

$$(4.5) \quad T \leq \frac{2}{b} \log \frac{1}{\delta}.$$

Consider the case with $\varepsilon > 0$. For each $i = 1, \dots, I$, the stable/unstable manifold of $(x_i^{c,\varepsilon}, 0)$ grows from $(x_i^{c,\varepsilon}, 0)$ and has the first contact with either the line $p = 0$ at $(x^*, 0)$, where x^* is not necessarily certain $(x_j^{c,\varepsilon}, 0)$, or the line $x = x_i^{c,\varepsilon}$ at $(x_i^{c,\varepsilon}, p^*)$ after passing over whole \mathbb{T} , where p^* is not necessarily 0. Let $(x(s), p(s))$ be any trajectory on the part of the stable/unstable manifold connecting $(x_i^{c,\varepsilon}, 0)$ and $(x^*, 0)$ or connecting $(x_i^{c,\varepsilon}, 0)$ and $(x_i^{c,\varepsilon}, p^*)$. Consider δ -cylinders $[x_j^{c,\varepsilon} - \delta, x_j^{c,\varepsilon} + \delta] \times \mathbb{R}$, $j = 1, \dots, I$ and the δ -cylinder $[x^* - \delta, x^* + \delta] \times \mathbb{R}$. Since the local stable/unstable manifolds of $(x_i^{c,\varepsilon}, 0)$ are close to those of $(x_i, 0)$ for sufficiently small $\varepsilon > 0$, we see, with a smaller $b > 0$ if necessary, that the above trajectory $(x(s), p(s))$ takes at most time $T > 0$ estimated as (4.5) to first touch one of the above cylinders after leaving a point outside of any of the cylinders.

Theorem 4.3. *Let $\delta(\varepsilon) > 0$ be any function tending to $0+$ with $\delta(\varepsilon) \geq \alpha\varepsilon$ as $\varepsilon \rightarrow 0+$, where α is a constant. Then, for each $c \in (c_-, c_+)$, the limit $v^* = \lim_{\varepsilon \rightarrow 0+} v^\varepsilon$ satisfies*

$$|v^\varepsilon(x) - v^*(x)| \leq \beta(\varepsilon |\log \delta(\varepsilon)|^2 + \delta(\varepsilon)) \quad \text{on } \mathbb{T} \quad \text{as } \varepsilon \rightarrow 0+,$$

where $\beta > 0$ is a constant independent of c, ε .

Remark 4.4. *For each $\nu \in (0, 1)$, we may take $\delta(\varepsilon) = \varepsilon^\nu$ obtaining $|v^\varepsilon(x) - v^*(x)| \leq \tilde{\beta}\varepsilon^\nu$ as $\varepsilon \rightarrow 0+$.*

Proof of Theorem 4.3. We compare the two value functions (2.1) and (3.1), where $h(c) = 0$ and $L(x, \xi) = \frac{1}{2}\xi^2 + F(x)$. In what follows, $\alpha_i > 0$ are some constants independent of ε and c . For any $\varepsilon \in (0, \varepsilon_0)$ with a sufficiently small $\varepsilon_0 > 0$, it holds that $|x_i^{c,\varepsilon} - x_i| < \alpha_1\varepsilon$ for all i and c . Take a number $\delta \geq \alpha_1\varepsilon$.

First we deal with the case of $c \in [c_-^*, c_+^*]$, where $\mathcal{M}_\alpha^\varepsilon(c) = \Gamma^{c,\varepsilon}$ and $|v^\varepsilon(x)| \leq \alpha_2\varepsilon$ on $\mathcal{M}_\alpha^\varepsilon(c)$ for $\varepsilon \rightarrow 0+$. Hence, $v(x) = 0$ on $\mathcal{M}_\alpha(c) = \{x_1, \dots, x_I\}$. Since v and v^ε are Lipschitz, we see that

$$(4.6) \quad |v^\varepsilon(x) - v^*(x)| \leq \alpha_2\delta \quad \text{on } [x_i - 2\delta, x_i + 2\delta] \quad \text{for } i = 1, \dots, I.$$

Suppose that the unstable manifold growing from $(x_i^{c,\varepsilon}, 0)$ first touches the line $p = 0$ away from any $(x_j^{c,\varepsilon}, 0)$, where the contact point (there are two such contact points at most for each i) is denoted by $(x^*, 0)$, and contains a point of graph $(c + v_x^\varepsilon)$ in the δ -cylinder $(x^* - \delta, x^* + \delta) \times \mathbb{R}$. Then, we define the interval $K_i^{c,\varepsilon}(x^*)$ as $K_i^{c,\varepsilon}(x^*) :=$

$(x^* - \delta, x^* + \delta)$. Let $K^{c,\varepsilon}$ denote the union of all such interval $K_i^{c,\varepsilon}(x^*)$ over $i = 1, \dots, I$, where if there is no such $K_i^{c,\varepsilon}(x^*)$, we define $K^{c,\varepsilon} := \emptyset$. Let x be an arbitrary point of $\mathbb{T} \setminus (\cup_{1 \leq i \leq I} [x_i - 2\delta, x_i + 2\delta] \cup K^{c,\varepsilon})$ and let $\gamma^*, \gamma^\varepsilon$ be a minimizing curve for $v(x), v^\varepsilon(x)$, respectively. Then $\gamma^*, \gamma^\varepsilon$ reaches one of $[x_i - 2\delta, x_i + 2\delta]$, $i = 1, \dots, I$ at most within time $T > 0$ estimated as (4.4). Hence, we have with (4.6),

$$\begin{aligned}
v^\varepsilon(x) - v^*(x) &\leq \int_{-T}^0 (e^{\varepsilon s} - 1) \{L(\gamma^*(s), \gamma^{*\prime}(s)) - c\gamma^{*\prime}(s)\} ds \\
&\quad + (e^{-\varepsilon T} - 1)v^\varepsilon(\gamma^*(-T)) + v^\varepsilon(\gamma^*(-T)) - v^*(\gamma^*(-T)) \\
&\leq \alpha_3(\varepsilon T^2 + \varepsilon T + \delta), \\
v^\varepsilon(x) - v^*(x) &\geq \int_{-T}^0 (e^{\varepsilon s} - 1) \{L(\gamma^\varepsilon(s), \gamma^{\varepsilon\prime}(s)) - c\gamma^{\varepsilon\prime}(s)\} ds \\
&\quad + (e^{-\varepsilon T} - 1)v^\varepsilon(\gamma^\varepsilon(-T)) + v^\varepsilon(\gamma^\varepsilon(-T)) - v^*(\gamma^\varepsilon(-T)) \\
&\geq -\alpha_3(\varepsilon T^2 + \varepsilon T + \delta).
\end{aligned}$$

Therefore, we obtain

$$|v^\varepsilon(x) - v^*(x)| \leq \alpha_3(\varepsilon T^2 + \varepsilon T + \delta) \quad \text{on } \mathbb{T} \setminus K^{c,\varepsilon}.$$

Let x be an arbitrary point of $K^{c,\varepsilon}$. Then, there is $\tilde{x} \neq K^{c,\varepsilon}$ such that $|x - \tilde{x}| \leq \delta$ and

$$\begin{aligned}
|v^\varepsilon(x) - v^*(x)| &= |v^\varepsilon(x) - v^*(x) - (v^\varepsilon(\tilde{x}) - v^*(\tilde{x})) + (v^\varepsilon(\tilde{x}) - v^*(\tilde{x}))| \\
&\leq |v^\varepsilon(x) - v^\varepsilon(\tilde{x})| + |v^*(x) - v^*(\tilde{x})| + |v^\varepsilon(\tilde{x}) - v^*(\tilde{x})| \\
&\leq \alpha_4\delta + \alpha_3(\varepsilon T^2 + \varepsilon T + \delta).
\end{aligned}$$

Thus, we conclude that

$$|v^\varepsilon(x) - v^*(x)| \leq \alpha_5(\varepsilon T^2 + \delta) \leq \alpha_6(\varepsilon |\log \delta|^2 + \delta) \quad \text{on } \mathbb{T},$$

where we may take any $\delta = \delta(\varepsilon)$ which tends to 0 with $\delta(\varepsilon) \geq \alpha_1\varepsilon$.

Now we deal with the general case. Let $J_i \subset (0, \varepsilon_0)$, $i = 1, \dots, I$ be such that if $\varepsilon \in J_i$ we have $x_i^{\varepsilon,c} \in \mathcal{M}_\alpha^\varepsilon(c)$. If $\inf J_i = 0$, we have $v^*(x_i) = 0$ due to (iii) of Theorem 4.1. Let $\varepsilon_* > 0$ be the minimum value of $\{\inf J_i \mid i \text{ is such that } \inf J_i > 0\}$. If there is no such i , re-define ε_* as $\varepsilon_* := \varepsilon_0$. Note that, if $x_i^{\varepsilon,c} \in \mathcal{M}_\alpha^\varepsilon(c)$ for some $\varepsilon \in (0, \varepsilon_*)$, we have $v^*(x_i) = 0$. For each sufficiently small $\varepsilon \in (0, \varepsilon_*)$, there exists at least one i for which $x_i^{\varepsilon,c} \in \mathcal{M}_\alpha^\varepsilon(c)$ and $v^*(x_i) = 0$. Suppose that $x_{i+1}^{\varepsilon,c} \notin \mathcal{M}_\alpha^\varepsilon(c)$. Then, one of the following two cases hold to be true:

- (i) $c + v_x^\varepsilon(x_{i+1}^{\varepsilon,c})$ is negative and hence $c + v_x^\varepsilon(x)$ is negative for all $x \in [x_{i+1}^{\varepsilon,c}, x_{i+2}^{\varepsilon,c})$,
- (ii) $c + v_x^\varepsilon(x_{i+1}^{\varepsilon,c})$ is positive and hence $c + v_x^\varepsilon(x)$ is positive for all $x \in (x_i^{\varepsilon,c}, x_{i+1}^{\varepsilon,c})$,

where we carefully note that discontinuity of $c + v_x^\varepsilon$ is allowed to be jump down only, and so is discontinuity of $c + v_x^*$. Since $c + v_x^\varepsilon$ converges to $c + v_x^*$ pointwise a.e. and hence uniformly except an arbitrarily small neighborhood of points of discontinuity as $\varepsilon \rightarrow 0+$, the sign of $c + v_x^\varepsilon$ is identical with that of $c + v_x^*$ except a small neighborhood of points of discontinuity and zero points of $c + v_x^\varepsilon$ and $c + v_x^*$.

If (i) is the case: We have $x_{i+k}^{c,\varepsilon} \in \mathcal{M}_\alpha^\varepsilon(c)$ such that $\text{graph}(c+v_x^\varepsilon)$ within $[x_{i+1}^{c,\varepsilon}, x_{i+k}^{c,\varepsilon}]$ coincides with the unstable manifold growing from $(x_{i+k}^{c,\varepsilon}, 0)$, which implies that $v^*(x_{i+k}) = 0$ and $\text{graph}(c+v_x^*)$ within $[x_{i+1}, x_{i+k}]$ coincides with the stable/unstable manifolds lying on $p \leq 0$ without any discontinuity. Hence, we obtain by the same procedure as the above,

$$|v^\varepsilon(x) - v^*(x)| \leq \alpha_7(\varepsilon T^2 + \varepsilon T + \delta) \quad \text{on } [x_{i+k-1} + 2\delta, x_{i+k} + 2\delta].$$

Then we obtain by Lipschitz continuity of v^ε and v^* ,

$$|v^\varepsilon(x) - v^*(x)| \leq \alpha_7(\varepsilon T^2 + \varepsilon T + \delta) + \alpha_8\delta \quad \text{on } [x_{i+k-1} - 2\delta, x_{i+k-1} + 2\delta],$$

which yields the estimate within $[x_{i+k-2} + 2\delta, x_{i+k-1} - 2\delta]$. We may repeat this argument with the above treatment in $[x_i, x_{i+1} - 2\delta] \cap K^{c,\varepsilon}$ if $K^{c,\varepsilon}$ is non-empty, and obtain

$$|v^\varepsilon(x) - v^*(x)| \leq \alpha_9(\varepsilon T^2 + \varepsilon T + \delta) \quad \text{on } [x_i, x_{i+k}].$$

If (ii) is the case: We have $x_{i+k}^{c,\varepsilon} \in \mathcal{M}_\alpha^\varepsilon(c)$ for which $\text{graph}(c+v_x^\varepsilon)$ within $[x_i^{c,\varepsilon}, \bar{x}^{c,\varepsilon}]$ with some $\bar{x}^{c,\varepsilon} \in (x_{i+k-1}^{c,\varepsilon}, x_{i+k}^{c,\varepsilon}]$ coincides with the unstable manifold growing from $(x_i^{c,\varepsilon}, 0)$ and, if $\bar{x}^{c,\varepsilon} \neq x_{i+k}^{c,\varepsilon}$, the graph has discontinuity at $\bar{x}^{c,\varepsilon}$ switching to the unstable manifold of $(x_{i+k}^{c,\varepsilon}, 0)$ lying on $p \leq 0$. This implies that $\text{graph}(c+v_x^*)$ within $[x_i, \bar{x}]$ with some $\bar{x} \in (x_{i+k-1}, x_{i+k}]$ coincides with the stable/unstable manifolds lying on $p \geq 0$ without any discontinuity and, if $\bar{x} \neq x_{i+k}$, the graph has discontinuity at \bar{x} switching to the unstable manifold of $(x_{i+k}, 0)$ lying on $p \leq 0$. Then we obtain with the same procedure as the above,

$$|v^\varepsilon(x) - v^*(x)| \leq \alpha_{10}(\varepsilon T^2 + \varepsilon T + \delta) \quad \text{on } [x_i, x_{i+1} + 2\delta],$$

and by repeating this argument with the same treatment in $[x_{i+k-1}, x_{i+k} - 2\delta] \cap K^{c,\varepsilon}$ if $K^{c,\varepsilon}$ is non-empty,

$$|v^\varepsilon(x) - v^*(x)| \leq \alpha_{11}(\varepsilon T^2 + \varepsilon T + \delta) \quad \text{on } [x_i, x_{i+k}].$$

In this way, we conclude $|v^\varepsilon(x) - v^*(x)| \leq \alpha_{12}(\varepsilon |\log \delta|^2 + \delta)$ on \mathbb{T} , where we may take any $\delta = \delta(\varepsilon)$ which tends to 0 with $\delta(\varepsilon) \geq \alpha_1 \varepsilon$. \square

4.2 Error estimate in (C2)

Suppose that a viscosity solution v of (HJ) admits a KAM n -torus. This means the following: $\text{graph}(c+v_x)$ is a smooth ϕ_H^s -invariant torus on which the Hamiltonian dynamics (2.4) is C^1 -conjugate to the linear flow on \mathbb{T}^n with a ν, η -Diophantine rotation vector $\omega \in \mathbb{R}^n$, i.e., there exists C^1 -embedding $\Phi = (\varphi, \psi) : \mathbb{T}^n \rightarrow \mathbb{T}^n \times \mathbb{R}^n$ such that

$$\begin{aligned} \Phi(\mathbb{T}^n) &= \text{graph}(c+v_x), \\ \phi_H^s(x_0, p_0) &= \Phi(\omega s + \theta_0) \text{ for all } (x_0, p_0) \in \text{graph}(c+v_x), \\ |\omega \cdot z| &\geq \nu(|z_1| + \dots + |z_n|)^{-\eta} \text{ for all } z \in \mathbb{Z}^n \setminus \{0\} \ (\nu > 0, \eta > n-1), \\ \theta_0 &\text{ is a constant depending on } (x_0, p_0). \end{aligned}$$

Note that in such a case, viscosity solutions v are unique up to constants. Since the linear flow with a Diophantine rotation vector is ergodic on \mathbb{T}^n , each minimizing curve

$\gamma^*(s)$ for $v(x_0)$, satisfying $\gamma^*(s) \equiv \varphi(\omega s + \theta_0)$, is also ergodic on \mathbb{T}^n . According to [9],[4], for each $\delta > 0$ the set

$$\mathcal{N}_\delta := \{\omega s + \theta_0 \mid -\frac{\beta_0}{\delta^\eta} \leq s \leq 0\} \quad (\beta_0 > 0 \text{ is a constant})$$

is δ -dense in \mathbb{T}^n , i.e.,

$$\bigcup_{\zeta \in \mathcal{N}_\delta} B_\delta(\zeta) = \mathbb{T}^n,$$

where $B_\delta(\zeta)$ is the δ -ball of around ζ . Hence, along only one minimizing curve $\gamma^*(s)$ within $-\frac{\beta_0}{\delta^\eta} \leq s \leq 0$, we obtain the whole information on \mathbb{T}^n in “ δ -accuracy”. This strategy is first used by Bessi [3] in the vanishing viscosity method, and by Soga [23] in the finite difference approximation.

Theorem 4.5. *Let v be a viscosity solution of (1.1) corresponding to a KAM torus with a ν, η -Diophantine rotation vector. Let v^ε be the unique viscosity solution of (1.2). Then, adding a constant which may depend on ε to v if necessary, we have for $\varepsilon \rightarrow 0$,*

$$\|v^\varepsilon - v\|_{C^0(\mathbb{T}^n)} \leq \beta \varepsilon^{\frac{1}{1+2\eta}},$$

where $\beta > 0$ is a constant independent of ε .

Remark 4.6. *There exists $v^* = \lim_{\varepsilon \rightarrow 0+} v^\varepsilon$, which is identical with v up to constant. The theorem does not necessarily mean that the error between v^ε and v^* is estimated by $\beta \varepsilon^{\frac{1}{1+2\eta}}$. The possible added constant also goes to 0 as $\varepsilon \rightarrow 0$.*

Proof of Theorem 4.5. In what follows, β_i are some positive constants. Adding a constant if necessary, we have

$$(4.7) \quad v^\varepsilon(x) - v(x) \leq 0 \text{ in } \mathbb{T}^n, \quad v^\varepsilon(x^*) - v(x^*) = 0 \text{ for some } x^* = x^*(\varepsilon) \in \mathbb{T}^n.$$

Let γ^* be a minimizing curve for $v(x^*)$. We see that

$$\begin{aligned} v(x^*) &= \int_{-\tau}^0 (L(\gamma^*(s), \gamma^{*'}(s)) - c \cdot \gamma^{*'}(s) + h(c)) ds + v(\gamma^*(-\tau)), \\ v^\varepsilon(x^*) &\leq \int_{-\tau}^0 e^{\varepsilon s} (L(\gamma^*(s), \gamma^{*'}(s)) - c \cdot \gamma^{*'}(s) + h(c)) ds + e^{-\varepsilon \tau} v^\varepsilon(\gamma^*(-\tau)). \end{aligned}$$

Subtracting $v(x^*)$ from $v^\varepsilon(x^*)$ yields

$$\begin{aligned} &v^\varepsilon(x^*) - v(x^*) \\ &\leq \int_{-\tau}^0 (e^{\varepsilon s} - 1)(L(\gamma^*(s), \gamma^{*'}(s)) - c \cdot \gamma^{*'}(s) + h(c)) ds + e^{-\varepsilon \tau} v^\varepsilon(\gamma^*(-\tau)) - v(\gamma^*(-\tau)) \\ &= v^\varepsilon(\gamma^*(-\tau)) - v(\gamma^*(-\tau)) \\ &\quad + \underbrace{\int_{-\tau}^0 (e^{\varepsilon s} - 1)(L(\gamma^*(s), \gamma^{*'}(s)) - c \cdot \gamma^{*'}(s) + h(c)) ds + (e^{-\varepsilon \tau} - 1)v^\varepsilon(\gamma^*(-\tau))}_{(\#)}. \end{aligned}$$

Since $|L(\gamma^*(s), \gamma^{*'}(s)) - c \cdot \gamma^{*'}(s) + h(c)| \leq \beta_1$ for all $s \in \mathbb{R}$, $|v^\varepsilon| \leq \beta_2$ independently of ε and $e^{\varepsilon s} \leq 1 + \varepsilon s + \frac{(\varepsilon s)^2}{2}$, $e^{\varepsilon s} \geq 1 + \varepsilon s$ for all $s \leq 0$, we obtain

$$|(\#)| \leq \beta_1 \int_{-\tau}^0 (1 - e^{\varepsilon s}) ds + \beta_2 (1 - e^{-\varepsilon \tau}) \leq \beta_1 \frac{\varepsilon \tau^2}{2} + \beta_2 \varepsilon \tau \quad \text{for any } \tau > 0.$$

By (4.7), we obtain $-(\sharp) \leq v^\varepsilon(\gamma^*(-\tau)) - v(\gamma^*(-\tau)) \leq 0$, i.e.,

$$|v^\varepsilon(\gamma^*(-\tau)) - v(\gamma^*(-\tau))| \leq \max\{\beta_3\varepsilon\tau^2, \beta_3\varepsilon\tau\} \quad \text{for any } \tau \geq 0.$$

Note that $\Phi = (\varphi, \psi)$ is an embedding and hence $\varphi : \mathbb{T}^n \rightarrow \mathbb{T}^n$ is diffeomorphic. Let $x \in \mathbb{T}^n$ be an arbitrary point. Set $\zeta_x := \varphi^{-1}(x)$. Since \mathcal{N}_δ is δ -dense in \mathbb{T}^n , we have $\zeta \in \mathcal{N}_\delta$ such that $\zeta_x \in B_\delta(\zeta)$. For this ζ , there exists $s \in [-\frac{\beta_0}{\delta\eta}, 0]$ such that

$$\zeta = \omega s + \theta_0 = \varphi^{-1}(\gamma^*(s)).$$

Therefore, we see that

$$\begin{aligned} & |v^\varepsilon(x) - v(x)| \\ & \leq |v^\varepsilon(x) - v^\varepsilon(\gamma^*(s))| + |v^\varepsilon(\gamma^*(s)) - v(\gamma^*(s))| + |v(\gamma^*(s)) - v(x)| \\ & = |v^\varepsilon(\gamma^*(s)) - v(\gamma^*(s))| + |v^\varepsilon(\varphi(\zeta_x)) - v^\varepsilon(\varphi(\zeta))| + |v(\varphi(\zeta_x)) - v(\varphi(\zeta))| \\ & \leq |v^\varepsilon(\gamma^*(s)) - v(\gamma^*(s))| + \beta_4|\zeta_x - \zeta| \\ & \leq \max\{\beta_5(\frac{\varepsilon}{\delta^{2\eta}} + \delta), \beta_5(\frac{\varepsilon}{\delta\eta} + \delta)\} = \beta_5(\frac{\varepsilon}{\delta^{2\eta}} + \delta). \end{aligned}$$

Taking $\delta = \varepsilon^{\frac{1}{1+2\eta}}$, we obtain the conclusion. \square

References

- [1] N. Anantharaman, On the zero-temperature or vanishing viscosity limit for certain Markov processes arising from Lagrangian dynamics, J. Eur. Math. Soc. **6** (2004), no. 2, 207–276.
- [2] N. Anantharaman, R. Iturriaga, P. Padilla and H. Sanchez-Morgado, Physical solutions of the Hamilton-Jacobi equation, Discrete Contin. Dyn. Syst. Ser. B **5** (2005), No. 3, 513–528.
- [3] U. Bessi, Aubry-Mather theory and Hamilton-Jacobi equations, Comm. Math. Phys. **235** (2003), 495–511.
- [4] J. Bourgain, F. Golse and B. Wennberg, On the distribution of free path lengths for the periodic Lorentz gas, Commun. Math. Phys. **190** (1998), 491–508.
- [5] F. Camilli, I. Capuzzo Dolcetta, D. A. Gomes, Error estimates for the approximation of the effective Hamiltonian, Appl. Math. Optim. **57** (2008), No. 1, 30–57.
- [6] P. Cannarsa and C. Sinestrari, *Semiconcave functions, Hamilton-Jacobi equations and optimal control*, Birkhäuser (2004).
- [7] A. Davini, A. Fathi, R. Iturriaga and M. Zavidovique, Convergence of the solutions of the discounted equation, Inv. Math., published online: 26 January 2016 (DOI: 10.1007/s00222-016-0648-6).
- [8] A. Davini, A. Fathi, R. Iturriaga and M. Zavidovique, Convergence of the solutions of the discounted equation: the discrete case, Math. Z., published online: 20 May 2016 (DOI: 10.1007/s00209-016-1685-y)

- [9] H. S. Dumas, Ergodization rates for linear flow on the torus, *J. Dyn. Differ. Equ.* **3** (1991), 593-610.
- [10] W. E, Aubry-Mather theory and periodic solutions of the forced Burgers equation, *Comm. Pure Appl. Math.* **52** (1999), No. 7, 811-828.
- [11] L. C. Evans and D. Gomes, Effective Hamiltonians and averaging for Hamiltonian dynamics. I, *Arch. Ration. Mech. Anal.* **157** (2001), No. 1, 1-33.
- [12] A. Fathi, Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens, (French) [A weak KAM theorem and Mather's theory of Lagrangian systems] *C. R. Acad. Sci. Paris Sér. I Math.* **324** (1997), No. 9, 1043-1046.
- [13] A. Fathi, *Weak KAM theorem in Lagrangian dynamics*, Cambridge Univ. Pr. (2011).
- [14] D. A. Gomes, Duality principles for fully nonlinear elliptic equations, *Progr. Non-linear Differential Equations Appl.* (2005) **61**, 125-136.
- [15] D. A. Gomes, Generalized Mather problem and selection principles for viscosity solutions and Mather measures, *Adv. Calc. Var.* **1** (2008), No. 3, 291-307.
- [16] D. A. Gomes, H. Mitake and H. V. Tran, The Selection problem for discounted Hamilton-Jacobi equations: some non-convex cases, preprint (arXiv:1605.07532).
- [17] H. Ishii, H. Mitake, and H. V. Tran, The vanishing discount problem and viscosity Mather measures. Part 1: the problem on a torus, preprint (arXiv:1603.01051), to appear in *Journal de Mathématiques Pures et Appliquées*.
- [18] H. R. Jauslin, H. O. Kreiss and J. Moser, On the forced Burgers equation with periodic boundary conditions, *Proc. Symp. Pure Math.* **65** (1999), 133-153.
- [19] N. Q. Le, H. Mitake and H. V. Tran, Dynamical and geometric aspects of Hamilton-Jacobi and linearized Monge-Ampere equations, to appear in *Lecture Notes in Mathematics*, Springer.
- [20] R. Mañé, Generic properties and problems of minimizing measures of Lagrangian systems, *Nonlinearity* **9** (1996), 273-310.
- [21] J. Mather, Action minimizing invariant measures for positive definite Lagrangian systems, *Math. Z.* **43** (1991), No. 2, 169-207.
- [22] H. Mitake and H. V. Tran, Selection problems for a discount degenerate viscous Hamilton-Jacobi equation, *Adv. Math.* **306** (2017), 684-703.
- [23] K. Soga, More on stochastic and variational approach to the Lax-Friedrichs scheme, *Math. Comp.* **85** (2016), 2161-2193.
- [24] K. Soga, Selection problems of \mathbb{Z}^2 -periodic entropy solutions and viscosity solutions, preprint (arXiv:1501.03594).